A Trajectory in First-Grade Children's Thinking About Variable and Variable Notation in Functional Relationships[1]

Maria Blanton
TERC
Cambridge, MA

Barbara M. Brizuela
Tufts University
Medford, MA

Katie Sawrey
Tufts University
Medford, MA

Angela Murphy Gardiner
TERC
Cambridge, MA

Ashley Newman-Owens
Tufts University
Medford, MA

Generalizing has received much attention—rightly so—as a core aspect of algebraic reasoning. Symbolizing a generalization, however, is no less significant (Kaput, 2008). The act of symbolizing derives its importance as the means by which one compresses multiple instances into the unitary form of a single statement that *symbolizes* the multiplicity (Kaput, Blanton, & Moreno, 2008). From this perspective, generalizing has been described as the "act of creating that symbolic object" (*ibid*, p. 20). To be clear, the symbolic object might be represented through either conventional or non-conventional forms. Indeed, it is possible—and even productive—to symbolize and reason symbolically through non-conventional forms such as one's natural language (Radford, 2011). However, few would disagree that algebraic reasoning ultimately involves reasoning with perhaps the most recognizable cultural artifact of algebra – the conventional symbol system based on variable notation (Kaput, 2008; Kline, 1972).[2]

The reconceptualization of algebra as a K–12 strand of thinking in the United States (US) (e.g., *Common Core State Standards Initiative* [CCSSI], 2010; National Council of Teachers of Mathematics [NCTM], 2000) raised the possibility of introducing algebraic concepts—including variable and variable notation—in grades earlier than previously thought possible. However, there are different perspectives on when and how to introduce variable notation. Part of the hesitation in introducing these concepts prior to middle grades lies, perhaps, in strict Piagetian interpretations of formal stages of development and the worry that premature formalism (Piaget, 1964) might lead students to meaningless actions on symbols, including those symbols used to represent variable quantities. Instead, some scholars emphasize that younger students should use those non-conventional systems that are already available to them—particularly natural language and drawings—as a means of symbolizing (Radford, 2011; Resnick, 1982). But as Carraher, Schliemann, and Schwartz (2008) insightfully observe, if we applied the argument that symbols should be introduced only when students know their meaning to first language learning, "adults would never speak to newborns on the grounds that infants do not already know what the words mean" (p. 237).

Even so, the concern of introducing premature formalisms gains some traction from the well-documented struggles *adolescents* have with the concept of variable and the use of variable notation (e.g., Knuth et al., 2011; Küchemann, 1981). One might extrapolate from this that these difficulties would only intensify with younger children. While we acknowledge this concern, recent research raises questions as to whether the struggles adolescents face are as tightly connected to age as previously believed. Some researchers have pointed out that, by middle grades, students have already developed particular ways of thinking about letters in contexts that are far removed from algebra—and even mathematics. Yet, they are then expected to build a new understanding of letters as a way to notate variable quantities in algebra (e.g., Braddon, Hall, & Taylor, 1993). This suggests that difficulties with variable notation may be less about age and the premature use of formalisms and more about conflicts the notation creates with the experiences and understandings students already have with the use of literal symbols in other non-mathematical contexts (McNeil, Weinberg, Hattikudar, Stephens, Asquith, Knuth, & Alibali, 2010).[3] Furthermore, others have found that well-designed curricular materials for middle grades, along with good instruction, can actually help students develop healthy conceptions of literal symbols as representing variables, conceptions that do not necessarily exhibit commonly observed difficulties such as object/quantity confusion (Knuth, Alibali, McNeil, Weinberg, & Stephens, 2005; MacGregor & Stacey, 1997; McNeil et al., 2010).

In conjunction with this, there is a growing body of early algebra research that suggests students in elementary grades (that is, grades K–5) have some facility with using variable notation to symbolize mathematical generalizations and, indeed, may actually choose variable notation over more non-conventional forms to represent these generalizations (e.g., Brizuela, Blanton, Gardiner, Newman-Owens, & Sawrey, 2014; Brizuela, Blanton, Sawrey, Newman-Owens, & Gardiner, 2014; Carraher, Schliemann, Brizuela, & Earnest, 2006; Schliemann, Carraher, & Brizuela, 2007).. A recent study examining the impacts of an early algebra instructional intervention in third grade found that students showed significant improvement in their ability to symbolize variable quantities, algebraic expressions, and functional relationships using variable notation and to reason with symbolized quantities to solve linear equations (Blanton, Stephens, Knuth, Gardiner, Isler, & Kim, 2014). Moreover, when these students were asked to represent functional relationships, they were more likely to use variable notation than natural language. Elsewhere, studies have shown that children in elementary grades can represent arithmetic properties and relationships, as well as relationships between non-specified, continuous quantities (e.g., area, length), using variable notation and can reason with these representations in their symbolic forms (Carpenter, Franke, & Levi, 2003; Dougherty, 2008, 2011).

There is reason, then, to be optimistic that a more systematic study of children's understanding of variable and variable notation is warranted. This should not be interpreted to mean that non-conventional representations (e.g., natural language) do not also play an important role in children's activity of generalizing—we think they do. But we wonder if an over-emphasis on non-conventional symbol systems in the elementary grades—particularly to the exclusion of conventional systems—is merited and whether the difficulties with variable notation seen among adolescents who have had no introduction to variable notation in elementary grades might be ameliorated by experiences in earlier grades through which they could begin to build a rich and nuanced understanding of variable and variable notation.

To this end, we set out to examine how children in their early years of formal schooling (grades K–2) understood variable and variable notation within the particular context of functional relationships. We recognize that because of the different roles variable can play (e.g., variable as a varying unknown, variable as a fixed unknown, variable as number in a generalized pattern [see Blanton, Levi, Crites, & Dougherty, 2010]), the study reported here is only part of the story. That is, functional relationships are not the only place where variable and variable notation occur, and so a full story should ultimately examine children's understanding of these ideas across different contexts. However, functional thinking is viewed as an important pathway into early algebraic thinking (Carraher & Schliemann, 2007) *and* an organizing strand across the K–12 curriculum (Freudenthal, 1982; Hamley, 1934; Schwartz, 1990). Moreover, of the different bodies of work within early algebra research (e.g., generalized arithmetic), functional thinking has arguably received the least attention in the *lower* elementary grades. As such, it represents an area of great need in early algebra research, particularly for this population of students.

Finally, we chose to focus on lower elementary grades because we thought students' lack of formal schooling—and lack of experiences with literal symbols—might help us flesh out boundary points in how their thinking can emerge. In this, we intentionally took an approach in which we set aside developmental constraints that might impose limits on children's experiences with variable and variable notation and, instead, sought to understand *how* they made sense of ideas often viewed as beyond their realm of thought. The goal of the study reported here, then, is to contribute to our understanding of the cognitive foundations of young children's thinking about variable and variable notation within functional relationships. The study is framed around the following research question:

> What levels of understanding about variable and variable notation do young children exhibit as they explore functional relationships between two quantities?

Research Design

We used learning trajectory research (Simon, 1995) as the conceptual mechanism for systematically exploring the development of children's understanding of variable and variable notation in functional relationships. A learning trajectory characterizes increasingly sophisticated levels in students' thinking about a concept as they progress through a particular instructional sequence (Clements & Sarama, 2009; Baroody, Cibulskis, Lai & Li, 2004). As this suggests, a trajectory is tightly coupled with the instructional sequence that gives rise to it and, as such, represents one *possible* progression in thinking. Moreover, the levels of a trajectory do not reflect a linear progression in one's thinking. That is, as the data presented here illustrate, students may skip levels of thinking in a trajectory on a particular task or may revert to lower levels in a trajectory when the instructional or environmental context changes (Sarama & Clements, in press).

Methodologically, we employed design research (Cobb et al., 2003; Kelly, 2003) in order to empirically develop hypothesized levels of thinking and to refine instructional sequences and the hypothesized levels in response to progressions in students' thinking (Collins, Joseph, & Bielaczyc, 2004; Stevens, Shin, & Krajcik, 2009). Classroom teaching experiments [CTEs] (Romberg, 1992; Steffe & Thompson, 2000) are commonly used to identify learning trajectories because their coordination of instructional design with informal, ongoing analysis during the teaching experiment and retrospective formal analysis of data generated from the experiment allows for the empirical refinement of a trajectory (Lesh & Lehrer, 1999). As such, we used CTEs in conjunction with individual interviews to explore levels of sophistication in children's thinking as they participated in our instructional sequence. In what follows, we describe our participants, the design of our instructional sequence, the implementation of the study, and our analysis of data.

Participants

One hundred and fifteen students, approximately ages 5 to 7, in two classrooms at each of grades K–2 participated in the study. The six participating classrooms were from two elementary schools ("School A" and "School B") in the Northeastern part of the United States. The schools were selected because they represented distinct demographics (summarized in Table 1).

School	Non-White	English as a Second Language	Low Socio-economic Status	Total Participants by Grade
School A	8%	5%	20%	21(K); 18 (1); 15 (2)
School B	98.6%	33.9%	89.5%	18 (K); 22 (1); 21 (2)

Table 1. Demographics for the two participating schools.

Design of Our Instructional Sequence

We briefly describe here the design of the instructional sequence used in our study. (For a more detailed description, see Blanton, Brizuela, Gardiner, Sawrey, & Newman-Owens, 2014.) We designed the lessons constituting the instructional sequence based on tasks that had been used productively in previous empirical research on children's functional thinking, particularly in grades 2–5 where the greater part of research on children's understanding of functional relationships has occurred (e.g., Blanton et al., 2014; Cooper & Warren, 2011; Moss, Beatty, Shillolo, & Barkins, 2008; Moss & McNab, 2011). Lessons were designed to explore functions of the type $y = mx$ and $y = x + b$. (A function of the type $y = mx + b$, $b \neq 0$, was not addressed until the post-interviews, discussed below.)[4] Because of the ages of the participants, constant values for m and b corresponded to values that were arithmetically meaningful for students of these ages. For example, functional relationships that might formally be represented as $y = 2x$ were used to leverage children's experiences with "counting by two's" and "doubling." Overall, tasks used in the instructional sequence were designed to have a common set of goals:

- Generate a set of co-varying data from a given problem situation;
- Organize the data in a function table and examine their meaning in terms of the problem situation;
- Explore and describe any relationships noticed in the data;
- Predict near and far function values;
- Explore unknown quantities and represent them through variable notation;
- Generalize a relationship between the two co-varying quantities and represent the relationship in words and variable notation;
- Describe the meaning for any variable notation, used to represent a quantity or relationship, in terms of the problem context.

Prior to the implementation of the study, a subset of tasks (i.e., at least one for each function type) was piloted in grades K–1 through classroom instruction and individual interviews. Our goal was to better understand how students might make sense of the tasks, to alert us to potential areas of concern, to modify surface features such as wording of the tasks, and to design instructional components that might help us better engage students (e.g., whether and how to use manipulatives; how to balance group work and whole-class discussions). Students participating in our pilot work did not participate in the study reported here.

The project team met weekly to discuss observations from pilot work and finalize the lessons for the instructional sequence. The instructional sequence was then organized as a set of two, four-week instructional cycles. Each cycle included eight 30-40 minute lessons (16 lessons total), with two lessons taught during each week of instruction. Each lesson was based on one task (see Appendix for the task "Attached Tables" constituting one of the lessons). Lessons in the first cycle (Cycle 1) focused on functional relationships of the form $y = mx$. The first week of Cycle 1, however, focused on repeating and growing geometric patterning tasks as a way to bridge the transition to function tasks. We used this time to introduce students to making the independent variable explicit, although we did not use the terminology "independent variable" with students. Lessons in the second cycle (Cycle 2) focused on functional relationships of the form $y = x + b$.

Implementation: Classroom Teaching Experiments and Individual Interviews

As described earlier, we conducted CTEs in six classrooms, two at each of grades K–2. The lessons were taught by one member of the project team at each site and were each observed and videotaped by other team members. The project team met weekly to discuss observations from the lessons focusing on children's thinking about concepts addressed in the function tasks and to discuss possible revisions in subsequent tasks.

Semi-clinical, individual interviews were also conducted in order to more closely detail levels of sophistication in individual students' functional thinking over the course of the CTEs that we might not have been able to glean from the uninterrupted flow of classroom instruction. In particular, we conducted pre-interviews prior to Cycle 1, mid-interviews between Cycles 1 and 2, and post-interviews after Cycle 2 with ten students from each of grades K–2 (30 students total). Each interview lasted about 30 minutes and was videotaped. During the interviews, we asked students to solve a task similar to those used in the CTE lessons and to describe his or her thinking aloud. For the interviews, we initially selected students who were identified by their teachers as being in the upper 30% of the class in terms of their understanding of arithmetic and ability to talk about their thinking. While it is important to understand students' thinking across mixed abilities, our emphasis at this early stage of research was to select interviewees who could complete at least some portion of the tasks we had designed. We chose a conservative approach prior to the implementation of the study—when interviewees were selected—to better ensure that the implementation of the study would be successful. However, as we implemented our study, we selected other students that were not necessarily within the top 30% academically, but that were able to verbalize their thinking well during our lessons.

Table 2 summarizes the structure of the CTEs and individual interviews and provides the core focus for a selected task from each cycle (see Appendix for a condensed version of the task "Attached Tables" as an illustration of task design; see Blanton et al., 2014, for a complete list of all tasks used and their core focus.)

Event	Function Type	Focus of Selected Task
Pre-Interview	$y = mx$	"Dogs and Noses" Find a relationship between the number of dogs and the total number of noses on the dogs.
Cycle 1 4 weeks	$y = mx$	Sample Tasks: "Attached Tables": $y = x + x$ Find a relationship between the number of square desks and the number of people that can be seated at the desks if the desks are joined end-to-end, no one sits on the ends, and one person can sit on each of two sides of a desk. (Blanton et al., 2014)
Mid-Interview	$y = x + \mathrm{b}$	"Height with a Hat" Find a relationship between a person's height and their height when wearing a hat that is 1 foot tall.

Cycle 2 4 weeks	$y = x + \mathrm{b}$	Sample Task: “Candy Boxes”: $y = x + 1$ Find a relationship between the number of candies John and Mary each have, if John and Mary each have a box of candies with the same number of candies and, in addition, Mary has one more candy on top of her box (Carraher et al., 2008)
Post-Interview	$y = mx$ $y = mx + b$	“Growing Train” Find a relationship between the number of stops a train makes and the total number of cars it has if the train adds two cars at each stop. Find a relationship for the case when the engine is not counted in the number of cars, and when it is counted.

Table 2. Structure of the instructional sequence and focus for sample tasks.

Data Analysis

The primary data source for the study reported here were videotaped individual interviews and students written work produced during the interviews. The videotaped classroom lessons from the CTEs served as a secondary source for triangulating findings about children’s thinking from the interviews. We used a grounded theory approach (Strauss & Corbin, 1990) whereby data analysis occurred in conjunction with the development of a (local) theory regarding a trajectory in children’s thinking about variable and variable notation as they progressed through our instructional sequence.

Retrospective data analysis occurred in two distinct phases. We began with an informal pre-analysis in which we collectively viewed selected videotapes to share initial insights about students’ thinking and identify possible boundary points (e.g., regarding how students reacted to the introduction of variable notation during the pre-interview). Our intent at this informal stage was not to identify potential codes or levels, but to elicit early reactions to students’ thinking about the tasks. We also used this phase to narrow our focus to first-grade students (or, six-year-old children) as a way to reduce our data set.

During formal data analysis, one member of the project team transcribed pre-, mid-, and post-interview video data for one first grade student and flagged any instances in the data related to variable or variable notation. In particular, data were flagged based on focal questions identified in the pre-analysis intended to be inclusive of incidents involving variable or variable notation: (1) How do students *represent* function rules?; (2) What representations are they using and how are they using these representations?; and—more specific to variable concepts—(3) How are they using variable and what understandings do they exhibit about variable? Flagged data were then analyzed line-by-line, and theoretical memos (Glaser, 1998) were constructed to characterize the incidents in which students used variable notation or discussed their understanding of variable quantities, their notation of these quantities, or what the notation represented. The memos consisted of descriptions of students' thinking, a possible interpretation of that thinking, and associated transcripts on which the descriptions were based.

Pre-, mid-, and post-interview data from another first-grade student were analyzed in a similar fashion and results from both analyses were compared to refine the memos. Memos were then sorted (Glaser, 2002) according to similar types of thinking and given a preliminary descriptive code that reflected the thinking within the memo group. The categories were organized based on their level of sophistication in children's thinking and an analysis of the mathematics to be learned (Battista, 2004). A preliminary learning trajectory was then hypothesized. The full project team then collectively analyzed the hypothesized trajectory and supporting memos and transcript data for the two students, and refined the trajectory to reflect the team's negotiated thinking about the levels. A second member of the team then independently reviewed the modified trajectory using a broader data set and, again, the full project team reviewed any subsequent revisions to the trajectory until agreement was reached.

Finally, using the revised trajectory, two other members of the project team independently coded the full set of first-grade interview data. Coding was based on the expectation that the levels of the trajectory would be refined as our analysis of a broader data set helped clarify earlier analyses of the data. To facilitate this stage of analysis, the remaining transcripts from interview video data were reviewed and flagged when any conversation addressed variable or variable notation. The flagged units of conversation were then coded according to the level or levels of thinking exhibited. Independent coding decisions were compared to determine inter-rater reliability. The two coders discussed any discrepancies in the coding in order to resolve the disagreement. If a discrepancy was not resolved, the full project team reviewed the transcripts and codes to reach a resolution (which might include that the data were inconclusive). The constant comparison and refinement of levels of the trajectory continued until no new findings emerged in the data.

Results

We describe here the main result of our study—a trajectory in children's thinking about variable and variable notation. In the levels presented here, there are several points to keep in mind. First, levels in a trajectory are not only intended to convey students' conceptualizations, but also what students "can and cannot do" (Battista, 2004, p. 187). Moreover, they should capture students' reasoning at the start of instruction as well as their formal mathematical understanding targeted in the instructional sequence (*ibid*).

We note again that the trajectory here is specifically linked to variable quantities associated with functional relationships. As such, we do not make any claims here about progressions in students' thinking about variable or variable notation in other contexts (for example, in equations with fixed unknowns or equations denoting arithmetic properties in which variable notation is used represent a number in a generalized pattern (Blanton et al., 2011)). Moreover, given a trajectory's tight connection to the instructional sequence from which it develops, it should be viewed as one *possible* progression in thinking (Baroody et al., Stevens et al., 2009).

Finally, in analyzing our data we found that children's thinking about variable was sometimes indistinguishable from their thinking about variable notation; at other times we observed distinct trajectories. Arguably, we could have treated these constructs independently and proposed a distinct trajectory for each. We chose instead to analyze how these constructs co-emerged in children's thinking. It was not the case that students' thinking about each of these constructs progressed to subsequent levels simultaneously. That is, there were times in which students' thinking about variable notation progressed, while their conception of a variable quantity seemed to remain unchanged. As such, in the discussion that follows, we have tried to clarify understandings about each of these constructs and the connections between them throughout the trajectory.

Level 1: Pre-Variable/ Pre-Symbolic

Children exhibiting characteristics of this level were "pre-variable" in that they did not recognize or did not yet perceive the presence of a variable in a mathematical context. Moreover, they were "pre-symbolic" in that they did not use any symbols—literal or non-literal—to represent these quantities. In other words, we would describe both the concept of variable and a symbolic system for notating it as outside of the child's perceptual field, where we take a perceptual field to include information (e.g., constructs) of which the child has some awareness. We do not take "pre-symbolic" to mean that the child has not had *any* experiences with symbols. To the contrary, by first grade—the grade level for the study reported here—children have had many experiences with literal symbols. However, we characterize children's thinking at this level as pre-symbolic in that they did not yet use literal symbols to *symbolize variable quantities.*

We note a particular symbiosis between pre-variable and pre-symbolic states. That is, if a child does not recognize a variable as a viable mathematical construct (i.e., is "pre-variable"), we would not expect him or her to use some set of inscriptions, such as literal symbols, to symbolize that construct. In other words, the child must perceive a variable before he or she has a need to symbolize it. At the same time, a child can have experiences with components of a symbolic system that can *potentially* serve the purpose of symbolization in a mathematical context, but for which such symbolizing *has not yet occurred*. As this suggests, the concept of variable and an understanding of a set of inscriptions that can potentially serve as part of a symbolic system to represent that variable can co-emerge in children's thinking. For example, a child can have experiences with the inscription that some cultures use to symbolize the operation of addition ('+'), but where these experiences are not related to addition in the child's thinking. (For example, a child might see and repeatedly press the symbol '+' on her mother's phone.) Similarly, he or she can begin to develop an understanding of the operation of addition by joining sets of objects, but without knowing that the inscription '+' is sometimes used to symbolize this operation. At some point, however, these experiences coalesce, and the child comes to understand that the operation of addition is symbolized by '+'. We are interested here in how that process occurs with variable and variable notation, that is, the particular nexus in which children begin to *use a notational system as a way to symbolize variable quantities*.

We observed pre-variable thinking in children's responses to problem situations before variable notation was introduced (that is, when children were also "pre-symbolic," as we use the term here). The following episode occurred during Levon's pre-interview, in which the interviewer explored with him the identity relationship between an unknown number of dogs and the number of noses on those dogs. As the excerpt opens, the interviewer has just asked Levon how many noses there would be for any number of dogs:

1 Levon: Three?

2 Interviewer: Okay. So, what if you didn't know how many dogs there were?

3 Levon: You have to guess.

4 Interviewer: You have to guess. Is there a way to show how many dogs there are if you don't know how many there are?

5 Levon: Um, yes.

6 Interviewer: How?

7 Levon: You could see.

When asked to consider a situation that involved an unknown number of dogs, Levon gave two approaches that we view as related and that suggest he did not yet perceive a variable as an object to be mathematized in the problem. First, when asked how many noses there would be for any number of dogs, he responded, "Three," then explained, "You have to guess." When asked further how he might represent how many dogs there were if he didn't know the number of dogs, he suggested that "you could see" (hence, count) the unknown number of dogs. In other words, Levon's approach for making sense of the situation was to guess or count the number of dogs (or noses) *in order to find a numerical value*. This is quite logical, however, given that his experiences—both in and out of school—had likely involved only operating on *known* quantities to find an unknown, but never mathematizing an unknown quantity, or variable. Because his experiences until this point had not included quantities whose values were unknown and *could not* be found (or did not need to be), it is reasonable that he did not perceive the quantity as an "unknown." Moreover, he had no logical necessity for constructing a symbolic system to represent an as-yet-unperceived unknown. As such, his thinking remained pre-symbolic.

Pre-variable and pre-symbolic thinking sometimes persisted after both problem scenarios with variable quantities and variable notation were introduced in our CTE. That is, for children whose thinking was characteristic of Level 1, the concepts of variable and variable notation had not yet sufficiently developed—even after their introduction—so as to merge into a view of literal symbols as a tool for representing variable quantities. Kaput, Blanton, and Moreno (2008) discuss how symbol and referent, taken in a broader sense than that used here, can be experienced as separate as children's understandings coalesce towards more conventional ways of thinking. This coalescence, which they characterize as a socially mediated process of symbolization that iteratively transforms children's thinking about symbol and referent, captures for us here the process of children coming to understand both variable *and* variable notation as a system for notating variable quantities. We would characterize this transformation for children in our study as being initiated in the social plane (Vygotsky, 1978) through our introduction in classroom instruction and interviews of problem scenarios that involved variable quantities or of literal symbols as a way to notate these quantities.

The following excerpt from Miah's mid-interview illustrates this persistence of pre-variable/pre-symbolic thinking. At this point, Miah has participated in a four-week instructional cycle where problem scenarios containing variable quantities and the use of literal symbols as a means to represent these quantities have been introduced in classroom instruction. In this excerpt, the task was to examine a relationship between a person's (unknown) height and his height when wearing a one-foot hat:

8 Interviewer: What if I said someone was *y* feet tall [interviewer writes *y* on the student's paper]? What do you think about that?

9 Miah: Hmm…. Just measure them [shrugs].

10 Interviewer: Okay. But could that [circles *y*] stand for any…

11 Miah: No.

12 Interviewer: …thing? No? What does it stand for?

13 Miah: Nothing.

14 Interviewer: Nothing. It doesn't stand for anything.

15 Miah: No.…

It might be argued that, in Levon's episode (lines 1–7), the interviewer's references to "how many" or "show" prompted Levon to count. However, Miah's thinking—consistent with Levon's—suggests that this was not necessarily the case. That is, Miah, too, objects to both the use of a letter (*y*) to symbolize someone's unknown height (as evidenced by her remark that *y* "doesn't stand for anything") and to the fact that someone's height could exist in an indeterminate state as an unknown ("just measure them"). Thus, while a literal symbol had been interjected by the interviewer as a way to represent a person's unknown height, Miah's thinking about the symbol at this point did not include that it might represent a variable (line 14), nor did she perceive an unknown height as a variable to be mathematized (line 9). That is, her thinking remains pre-variable and pre-symbolic, although she has now been in instructional contexts where variable and variable notation are explored.

For a child such as Levon or Miah whose thinking is characteristic of Level 1, a reasonable move when encountering a variable in either a written symbolic situation (e.g., a variable quantity or function rule notated symbolically) or a problem scenario conveyed orally (e.g., “Mary has some candies in her box, but we don’t know how many”) would be to find a numerical value for the unknown quantity by guessing or counting. Miah does this—in theory—by proposing that the person’s height, which the interviewer refers to as y, be measured. We suggest that Miah did not conceptualize the person’s height as a variable—that is, she did not perceive it as an unknown—so it is reasonable that, for her, y could not “stand for anything.” That is, the referent did not exist, so the letter could not symbolize.

Level 2: Pre-Variable/Letters as Labels or as Representing Objects

Children whose thinking was characteristic of Level 2 viewed a letter as a label or as representing an object, but not as representing a variable. In this, they tacitly accepted the use of a letter as representing *something other than a variable*. That is, the letter was used as *symbolic* notation, but not *variable* notation. In particular, they used literal symbols as a label or as a way to represent (tangible) objects, such as the person whose height was unknown. We infer from this that children at this level were still “pre-variable”, that is, they still did not perceive unknown quantities as objects to be mathematized. Thus, in the absence—to them—of a variable as a referent to be symbolized, they used literal symbols to represent other things, such as objects or characters suggested to them from the problem context. The most visible aspect of a problem situation in a function task is arguably the object for which some quantity is being measured (e.g., a person whose height is being measured), so it is reasonable that a young child would symbolize the object in the absence of perceiving the quantity.

The following excerpt, which occurred during Jada’s mid-interview, exhibits this thinking. The excerpt opens after the interviewer has reminded Jada that she (the interviewer) used y to represent “how tall someone is” and $y + 1$ to represent “how tall they are with the hat.”

16 Interviewer: Does that make sense or not?

17 Jada: Y plus....

18 Interviewer: It's y plus one 'cause one is [the height of] the hat. Does that make sense or no?

19 Jada: It does because, because, so the person y, let's just pretend their name is y.

20 Interviewer: Yeah.

21 Jada: Then the person's name is y. So, the hat is one foot tall, so it does make sense because the person's name is y and....

Jada pauses for about 13 seconds, then adds, "I would have to measure them." We maintain that Jada did not yet perceive the variable in the situation because she ultimately abandoned her search to explain why the representations suggested by the interviewer (y and $y + 1$) made sense *as a way to represent someone's unknown height*, stating instead, "I would have to measure them." In other words, Jada interprets the quantity in question—a person's height—as something to be found by measuring, not symbolized as an unknown. Moreover, it is clear from the interviewer's use of literal symbols that variable notation had been introduced to Jada (by the mid-interview, Jada has also participated in the first 4-week instructional cycle). Jada's assignment of y as the name of—or, label for—the person whose height is unknown ("let's just pretend their name is *y"*) even suggests that she had taken up the use of a literal symbol and recognized that it was intended to symbolize *something unknown*. However, what it seemed to symbolize for Jada was an unknown person's name, not a mathematical unknown.

This type of thinking is well documented among adolescents (e.g., Booth, 1988; McNeil, 2010). What is less understood is why it occurs in young children as well. Jada's thinking about variable and variable notation at this point is consistent with our earlier claim that it is counterintuitive for a child to accept the use of an inscription to symbolize an object that the child does not yet perceive. That is, we contend that Jada did not see *y* as symbolizing a variable quantity because she did not yet perceive the quantity. However, the person whose name is unknown *did* exist for Jada in some sense. That is, she could mentally visualize such a person and imagine letting this person's (unknown) name be denoted by the letter *y*. It is reasonable in her thinking that the person (i.e., object) was the likely referent to be symbolized. Thus, while she accepted the literal-symbolic notation—that is, the use of a letter to notate *something*—she did not take it up as *variable* notation or as a means to symbolically represent variable quantities, in part because she did not yet perceive a variable as an object that could exist indeterminately and that needed to be mathematized.

Level 3: Letters as Representing Variables with Fixed, Deterministic Values

Children's thinking about variable quantities at Level 3 reflected a fundamental shift from that of previous levels in that they now perceived a variable and saw a literal symbol as representing that variable—that is, symbolic notation now acted as variable notation. However, even though the functional context of the problems addressed in our study was based on the role of variable as a varying unknown, children at Level 3 viewed variable as an unknown that had a fixed, deterministic value that could be determined by an internal logic that they perceived, but that was external to their control. The process was external in that they could not—or saw themselves as no longer allowed to—find the value of the unknown through their own actions, such as randomly guessing or physically counting or measuring. Instead, they viewed the value of the unknown as deterministically linked to the letter used to symbolize the variable quantity and its ordinal position in the alphabet. That is, they now perceived a variable and the literal symbol used to represent it also provided the mechanism for finding its value. In this sense, we describe this "fixed" view of a variable quantity as deterministic because its value was bound by the symbol representing it. For example, if the variable quantity was symbolized with the letter *D*, then in the child's view its value *must have been* 4. In this, however, their understanding of variable notation was still emerging because they did not yet see that the choice of symbol was arbitrary and, thus, that its value was not deterministically linked to the choice of symbol. This type of thinking has been framed in the literature as a misconception with regards to students' (particularly, adolescents') understanding of variable (MacGregor & Stacey, 1997; McNeil, 2010). For young children, however, we view this as an emerging conception that reflects their attempts to make sense of a new notational system whose symbols have another inherent—and non-mathematical—meaning. Even before children begin formal schooling, they develop an understanding of an alphabetic system whose order and structure might be naturally be brought to bear on their construction of a new—but related—symbolic system.

The following episode, which illustrates this type of thinking, occurred in the pre-interview with Rebecca, when she was first introduced to the use of variable notation to represent a relationship between an unknown number of dogs and the number of noses on the dogs:

22 Interviewer: I'm going to use a letter. I'm going to use *W*. I don't know how many dogs I have in my back yard. I have *W* dogs. So how many noses do I have?

23 Rebecca: Ohm….(Rebecca looks down at her fingers and starts counting audibly).

24 Interviewer: Tell me what you're doing.

25 Rebecca: Counting the alphabet to see how many letters are, are lower, I mean lesser, that's how much.

26 Interviewer: OK, so go ahead. You do that.

27 Rebecca: (Sings the alphabet song and counts.) Wait, there's 24 letters in the alphabet?

28 Interviewer: There's 26 letters in the alphabet.

29 Rebecca: (Counts backwards from the letter Z.) Twenty-six, twenty-five, twenty-four, twenty-three, twenty-two, twenty-one. Twenty-one dogs.

30 Interviewer: OK, so you think I have 21 dogs because *W* is the twenty-first letter of the alphabet. Is that right?

31 Rebecca: Yeah.

First, we infer from the fact that Rebecca took up the interviewer's question regarding the number of dog *noses* that she now accepted that there could be an unknown number of *dogs* in the backyard as well as the convention, new to her, that the unknown number of dogs could be represented by the letter *W*. That is, she now perceived a variable in the problem situation *and* she accepted the use of a literal symbol as variable notation. However, as with children's thinking in previous levels, she still needed a mechanism for finding the (fixed) value of the variable quantity. She determined the number of dogs, symbolized by *W*, to be 21 because of the ordinal position of *W* in the alphabet. In essence, this "alphabet strategy" worked as a system for decoding the value of a variable. However, a subtle but important difference in this type of thinking in comparison to that of previous levels is that the mechanism for finding this value was contained within an "alpha-mathematical" system in which Rebecca worked and that was *external* to her own choices and actions such as guessing and measuring. That is, this new strategy for finding the value of the literal symbol was based on an ordering of letters that was an accepted, authoritative norm and not her own. In our view, this lent to Rebecca legitimacy to her method that the value of *W must have been* 21 because of the established ordering of the alphabet. In the limited experiential world of a first-grader—an alpha-centric world still defined in terms of relatively small numbers—this was a logical choice for how to make sense of *W*.

As the interview continued, Rebecca represented the number of dog noses by writing *W* in the second column of her function table, next to the *W* representing the number of dogs in the first column (see Figure 1):

32 Interviewer: Oh, what did you just do?

33 Rebecca: A *W* (indicating that she wrote *W* in the second column).

34 Interviewer: Why?

35 Rebecca: For that much dogs, and then over here I'm going to write how many numbers. Twenty-one (she writes '21' beside *W* in the second column).

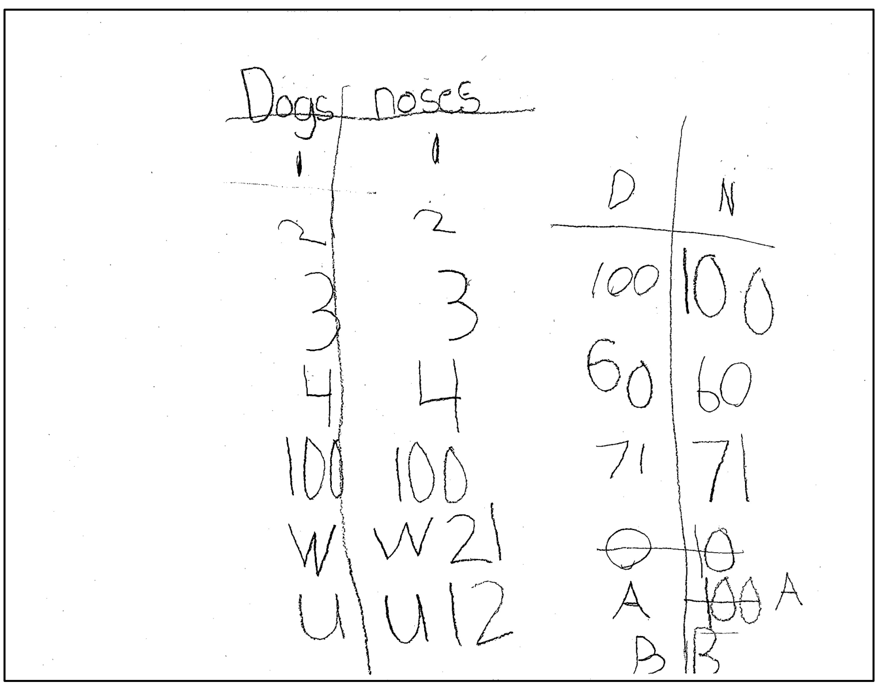


Figure 1. Rebecca's written work for the pre-interview task.

Rebecca's inscriptions suggest that she saw *W* as equivalent to 21—she could use either to represent the number of dogs. That she then used both *W* and 21 as representations of the number of dog noses suggests to us that she saw *W* as representing a quantity (a quantity whose value was 21). Moreover, that she constrained the value of *W* to 21 further supports that she did not, at this point, view W as a varying unknown. That is, she held a view of variable as a fixed unknown whose value is specific and could be found using established norms (that is, by following the order of the alphabet). In this sense, children at this level interpreted a variable that actually functioned as a varying, unknown quantity—as is the case for variables in functional relationships—*as a fixed unknown*. This view might be qualitatively distinct from a (correct) view of variable as a fixed, unknown in situations where it actually functions as such (e.g., the role of x in the equation $3 + x = 12$). That is, even though the value of the quantity here could actually be any number subject to satisfying the functional relationship, children whose thinking was characteristic of this level viewed that quantity as having a fixed, deterministic value.

Level 4: Letters as Representing Variables with Fixed but Arbitrarily Chosen Values. As with Level 3, children whose thinking was at Level 4 viewed a variable as a fixed unknown. However, an important distinction of Level 4 is that the unknown had a single, fixed value that could be *randomly* chosen. That is, their view of variable was no longer deterministic in that the letter symbolizing the variable determined its ("fixed") value. Instead, the child conceptualized a literal symbol as representing "any number," but conceived of "any number" as *any fixed number* that could be arbitrarily chosen and was not bound by a mathematical relationship or some choice that had an internal logic for the child (e.g., the "alphabet strategy"). The pre-interview with Rebecca, in which the task focused on exploring the identity relationship between an unknown number of dogs and the corresponding number of dog noses, illustrates this idea:

36 Interviewer: How many noses would I have if I had *U* number of dogs?

37 Rebecca: Twelve.

38 Interviewer: Ok, twelve. I could have twelve.

39 Rebecca: You could have twelve.

40 Interviewer: Why could I have twelve?

41 Rebecca: Because it's any number.

First, Rebecca's choice of "12" as the value of *U* was not based on the ordinal position of *U* in the alphabet. We note, however, that this type of thinking seems similar to that exhibited in Levels 1 and 2 in which children randomly chose (guessed) a value for the variable. We think a critical distinction is that at earlier levels, children did not perceive a variable, so their act of assigning a numerical value did not involve reasoning about either variable notation (as a symbol of) or a variable (as a referent to be symbolized or a quantity to be mathematized). At Level 4, however, students perceived a variable quantity, but understood that quantity to have a value that was fixed *and* arbitrarily chosen. For example, when the interviewer later asked Rebecca, "What if we had *P* noses, how many dogs would we have?", Rebecca responded, "Ohm…ninety-one." In this, Rebecca seemed to be taking up the notion of "any number" by choosing the number of dog noses—91—in a way that suggested she was free to choose any (fixed) number she wanted. She was not bound—as she was in Level 3 thinking— by a strategy that produced the only value that *P could possibly have been* (e.g., by using the "alphabet strategy"). In other words, the child's thinking at Level 4 seemed to be: (1) There is quantity whose value I do not know; (2) The value of that quantity can be any randomly chosen number; (3) Once the value is chosen, the quantity cannot have another value. We see this as a different view of "fixed" than is reflected, for example, in the value of x in the equation $3x + 6 = 12$. In an equation such as this, there is only a single, valid value for x ($x = 2$) and there is an underlying logic—algebraic syntax—that allows one to "find" the fixed value of x. In contrast, children's thinking at Level 4 allowed for the value of the unknown to be any number (that is, the value did not have to satisfy a mathematical constraint such as an equation), but once that number was chosen, it was fixed.

Level 5: Letters as Representing Variables that are Varying Unknowns

At Level 5, children conceptualized a variable in a functional context as a varying unknown and a literal symbol as representing a varying unknown. Recall that during Rebecca's pre-interview, she had initially said there would be 12 noses for *U* dogs (lines 36–41). She then wrote *U* in the second column of her function table to correspond to *U* dogs and wrote '12' beside the *U* in the second column (see Rebecca's written work in Figure 1). At that point in her (Level 4) thinking, Rebecca's notion of "any number" seemed to be that she could select *any* value and assign that value to the variable, but once the value for the variable was decided, it was fixed. We infer this partly because, when asked about the number of noses for *U* dogs, her response was simply "twelve" (line 37), in much the same way that she had assigned the value of 91 to *P*. If she had conceptualized the variable as a *varying* unknown, we would expect her to make some claim that reflected that 12 was a *possible* value of *U*, but that there could be other values for the variable. For example, a response such as "It could be 12, but it could also be 7 or 10" would reflect a conception of the value of the variable as varying. The conversation with Rebecca continued, and after a brief discussion, Rebecca and the interviewer had the following exchange:

42 Interviewer: Ok, so if I have *U* dogs, then I have *U* noses?

43 Rebecca: Yeah.

44 Interviewer: So, if I wasn't asking you for a number at all, we were just sticking to letters, we would say the relationship would be, if you have *U* dogs, you would have *U* noses?

45 Rebecca: Yeah.

46 Interviewer: How come?

47 Rebecca: Because it, uhm, because it could be any, uhm, numbers of the alphabet could be any number.

48 Interviewer: OK. Letters could be any number (clarifying that by “numbers” of the alphabet, Rebecca actually meant “letters” of the alphabet)?

49 Rebecca: It could be any number. Like twelve, two, one, zero.

50 Interviewer: Yeah, zero’s a number. OK, could *U* mean twelve?

51 Rebecca: Yeah.

52 Interviewer: And could *U* mean thirty-three?

53 Rebecca: Yeah.

54 Interviewer: So it could mean any number?

55 Rebecca: Yeah. Even a hundred.

56 Interviewer: Even a hundred. Wow. Is 100 the biggest number you know?

57 Rebecca: No.

58 Interviewer: Wow. Even bigger numbers? So, we’ll say that maybe *W* could have the value of 21 right now and maybe *U* could have the value of 12, but maybe it doesn’t, right? It could be any number?

59 Rebecca: (Nods yes.)

This (line 49) is the first point during the pre-interview at which Rebecca suggests that the value of the quantity could vary, indicating to us that her thinking is transitioning to a view of variable—in a functions context—as a varying quantity. We do not claim that her understanding of variable is robust at this point. Indeed, as we saw in subsequent interviews with Rebecca and other participants, and as others have observed elsewhere (e.g., Sarama & Clements, in press), students often revert to lower levels in a trajectory when faced with a new task or even benign alterations in their environment. We interpret Rebecca's thinking, instead, as the early seeds of a concept that can develop as she has more experiences with variable quantities and variable notation. It would be useful, then, to juxtapose this emergent view of variable and variable notation with children's conceptions later in the CTEs. The following episode occurred during the post-interview with Jackson, where the central focus was to explore a relationship between the number of stops a train makes and the number of train cars it has, assuming that the train picks up two cars at every stop and the engine is not counted. As the episode opens, the interviewer has asked Jackson to think about the case where the number of train stops is not known ("Suppose we didn't know how many stops the train made"):

60 Jackson: So it would be a number, *A*.

61 Interviewer: Tell me what *A* is representing.

62 Jackson: Like, the number of stops he was at, like 20, 40, 60 – any stop.

63 Interviewer: So I'm glad you said that. You said *A* could be 20, 40, 60, so *A* could be anything?

64 Jackson: Yeah.

65 Interviewer: Do we need to know what it is?

66 Jackson: No.

It seems clear that Jackson viewed the value of the variable quantity represented by *A* as varying. Perhaps an even more important characteristic of his thinking is that he *did not need to assign a value to the variable.* We see it as a significant marker of the sophistication in Jackson's thinking about variable that, for him, the variable could exist in an unresolved, indeterminate state in which its value was never found. That is, he does not have an intrinsic need to "find" the value of the variable, in the same way perhaps that a child who progresses beyond an operational understanding of the equal sign does not need to "compute" when encountering the equal sign.

Level 6: Letters Representing Variables as Mathematical Objects.

Children whose thinking was characteristic of Level 6 not only recognized a variable as a varying unknown and represented it with a literal symbol, they were also able to use variables (as well as constants and operations) to represent functional relationships (see Blanton et al., 2014, for a more detailed treatment of a trajectory in children's understanding of functional relationships). Furthermore, they were able to act on variables, represented with literal symbols, in novel situations as mathematical objects in their own right. The following selected excerpts depict conversations between Rebecca and the interviewer during the post interview, for which the task initially focused on finding the relationship between the number of stops a train made and the total number of cars, if the train picked up two cars at each stop. As the first excerpt opens, Rebecca has just completed a function table (a representation she chose to use without the interviewer's prompting) (see Figure 2). The interviewer asks her about a far value for the function:

67 Interviewer: So what if your train made 100 stops? How many cars would [it] have?

68 Rebecca: Two hundred?

69 Interviewer: OK. How did you get that?

70 Rebecca: Because you just double it. Because $1 + 1 = 2$, $2 + 2 = 4$, $3 + 3 = 6$, and $4 + 4 = 8$ (she writes these equations beside the corresponding values in the table).

71 Interviewer: Great. I like the way you did that. So, let's build on this. What if you didn't know how many stops your train made?

72 Rebecca: You could use a variable.

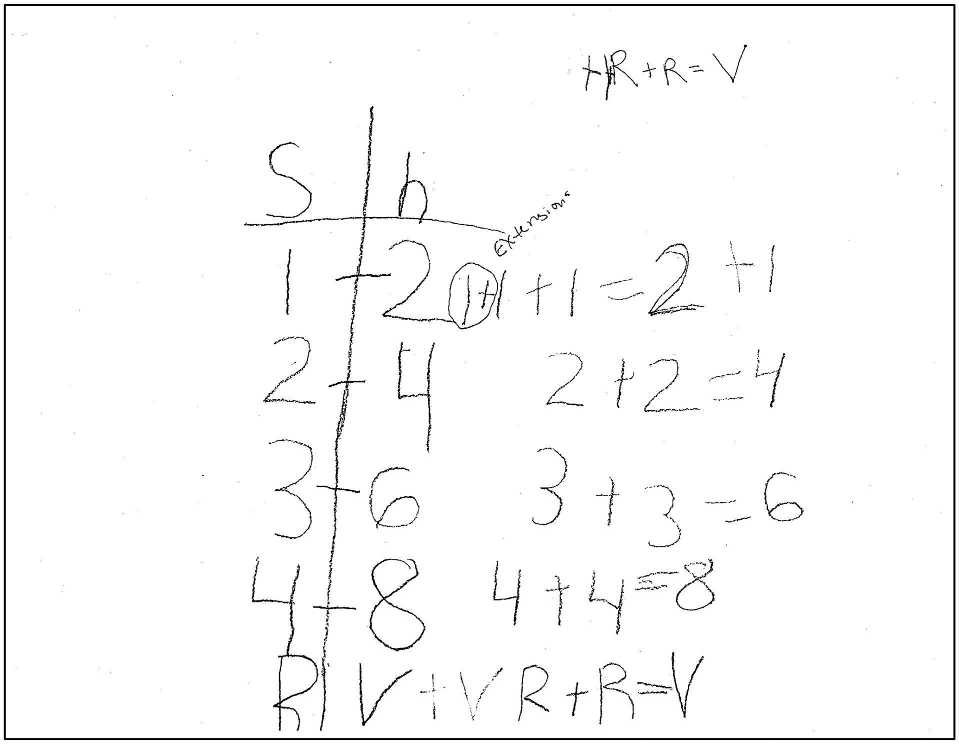


Figure 2. Rebecca's written work on the post-interview task.

When asked what letter she would use, she stated *R* because "that's the first letter of my name." Rebecca then used *V* to represent the total number of cars the train had after *R* stops and characterized the relationship between the two variables as $R + R = V$. Although the choice of *R* had personal meaning for Rebecca, the choice of *V* did not, suggesting to us that she understood the choice of letter to be arbitrary. To further probe Rebecca's thinking, the interviewer asked Rebecca about the meaning of the notation she used to represent the variables and the rule:

73 Interviewer: So, what does *R* represent here in your equation?

74 Rebecca: *R* represents how many stops the car, the train [pointing to the train] makes.

75 Interviewer: OK, and what does *V* represent?

76 Rebecca: How many carts (cars) he has.

These excerpts suggest that Rebecca both perceived a variable in the problem scenario and knew that it could be represented with a literal symbol. She understood the meaning of the notation she used within the problem context, and she could operate on generalized and symbolically notated quantities (R, V) as objects to produce a rule depicting a relationship ($R + R = V$). It is not trivial for a six-year-old to construct a representation in which it appears that letters are being "added" together and that that "result" is equivalent to another letter. Thus, while she might understand that R and V represented, respectively, the number of stops and the total number of train cars, a reasonable question is what sense did she make of the rule she constructed? The following excerpt, which occurred immediately after the episode in lines 73–76, shows Rebecca's interpretation of the symbols within the problem context:

77 Interviewer: Great. Let's suppose you take this home to your Mom…. So how would you explain this rule to your mom or friend in your own words?

78 Rebecca: Whatever number how many stops it made, if you doubled it, that's how many cars it would have.

Finally, not only was Rebecca able to articulate a rule to represent the relationship in the problem and interpret the meaning of the rule within the problem context, she also seemed to intuitively understand the boundaries of that rule and how perturbations in the problem situation might be reflected in its components. When asked earlier in the interview whether her rule, $R + R = V$, would always work, Rebecca suggested that if the engine *were* to be counted, it would change her rule. When later asked how her rule might change if she now counted the engine, Rebecca described that "you can just add a plus one" and symbolized the new relationship as "$+ 1\ R + R = V$" and later as "$+ 1 + R + R = V$." In other words, we maintain that the rule and its constituent parts—including variables—were objects themselves that she could transform or operate on to produce a (new) relationship.

Discussion

There are several observations we can make about the progression we observed in children's thinking about variable and variable notation as they participated in our instructional sequence. First, there was a subtle shift in their thinking about symbolic notation between Levels 1 and 2—a shift from not knowing to use or accepting the use of a literal symbol to symbolize an unknown quantity, to the interpretation of a literal symbol as representing something that was unknown but not inherently mathematical (such as the name of an unknown person). This misunderstanding of the referent—of what needed to be symbolized—seemed consistent with children's thinking about variable quantities across Levels 1 and 2, which held that unknown (mathematical) quantities did not or could not exist—that is, they were not yet perceived by the child—and, therefore, there was no intrinsic reason or capacity to symbolize them. At the same time, by Level 2 children recognized that a literal symbol might be used to represent *something*, although not yet a variable. That is, children were beginning to appropriate the use of literal symbols for symbolic notation, but not *variable* notation.

If we frame the trajectory in children's thinking from a process/object lens (Sfard, 1991), we might characterize the mental activity in Levels 1 and 2 as a process of interiorization in that concept formation had been initiated. However, children at these levels were still operating with (or invoking) familiar actions on known objects, such as counting or measuring quantities or proposing that these actions be performed. In this, they did not conceptualize the variable quantity in a problem situation as an unknown or indeterminate quantity.

Because children at Levels 1 and 2 did not seem to perceive a variable quantity in a problem situation, it seems natural that any reference to a quantity whose value was unknown (whether in symbolic or natural language) would be countered by their assignment of a numerical value to that unknown. In such cases, children either proposed or employed an external mechanism for finding the value (e.g., measuring someone's height, counting the number of dogs); alternatively, they sometimes randomly assigned a specific value to the unknown. In effect, producing the "fixed" value for the unknown quantity seemed to serve as a de facto argument that it did not exist. That is, they seemed to mathematize the problem in a way that treated variables as *known* quantities. Thus, in the absence of perceiving a variable, the child interpreted a situation involving a variable as a prompt to find a value by some familiar mechanism (e.g., counting, guessing).

In our view, this type of thinking is not unlike the misconception about the equal sign that develops in elementary grades through a prolonged study of arithmetic isolated from algebra (Carpenter, Franke, & Levi, 2003). That is, children in elementary grades learn to think operationally about the equal sign as a prompt to *compute* because of their extensive experiences with operating on *knowns* to find *unknowns* and with equations that reflect this type of action. In much the same way, we suggest that children also come to think "operationally" about unknown quantities as a prompt to *find* because they rarely—if ever—encounter problem situations where unknown quantities need to be mathematized.

Moreover, the thinking exhibited in Levels 1 and 2 suggests that the lack of perception of a variable may constrain or distort the development of a symbolic system to represent the quantity (in contrast to the situation where children might perceive a variable, but not yet have a means to symbolize it). Conventional wisdom has argued that children are not yet "ready" to use variable notation. We wonder, instead, if the problem lies not in children's readiness to use a symbolic system, but in their lack of experiences with mathematical situations that involve unknown quantities and that can motivate the need for children to construct such a system. That is, we wonder if the predominant focus in elementary grades mathematics on finding unknowns *given the knowns*—or, said another way, the lack of opportunities children are given to reason about unknown situations as indeterminates—has served as a didactical obstacle in the development of their mathematical thinking. If, instead, children were routinely provided with experiences to reason about and represent unknown quantities, would a symbolic system for notating these quantities arise more naturally as a logical necessity for representing their thinking? As McNeil et al (2010) suggest, "The process of generating representations may help students learn to use symbols meaningfully to represent unknown quantities (p. 632)."

The study reported here suggests that, just as children can learn to think relationally about the equal sign—a foundational early algebraic concept—so, too, can they learn to correctly mathematize and reason with unknown quantities. In Levels 3–5, children's thinking about variable and variable notation began to "condense" in that they began to perceive a variable quantity and represent this new construct via variable notation. An important characteristic of this shift in their thinking was the shift from a view of the value of a variable as being fixed and deterministic in Level 3, to a view of the value of the variable as being any *one*, fixed value from an implicit *range* of values in Level 4. This seemed to set the stage for a shift from a conception of variable as a "fixed unknown" (Levels 3 and 4) to that of a "varying unknown" (Level 5). In our view, what condensed across these levels was the conceptualization of the variable as a *varying unknown*—an important objective for the functional context in which this study occurred—as well as the use of literal symbols as a set of inscriptions that could serve as *variable notation*.

Finally, Level 6 reflected a reification of variable and variable notation in which children could mathematize unknown quantities and act on these quantities as objects in and of themselves, or even combine them with other symbols (operations, numerals) to represent functional relationships. Table 1 summarizes the levels in the trajectory proposed here and how we view their relationship to the process of reification in children's thinking about variable and variable notation.

Conceptual Process	Trajectory Levels	Characteristics
Interiorization Operating on known concepts	*Pre-Variable/Pre-Symbolic*	Does not recognize or perceive an unknown quantity in a mathematical situation; Does not use or accept the use of any symbolic notation (literal or non-literal) to represent a quantity.
	Pre-Variable/Letter as Label or as Representing Object	Tacitly accepts the use of a literal symbol to represent something not known, but not an unknown mathematical quantity (hence, an indication that child does not yet recognize a variable quantity).
Condensation	*Letter as Representing Variable with Fixed, Deterministic Value*	Recognizes a variable quantity and sees the literal symbol as representing the quantity; Value of the variable quantity is deterministic, often linked to the letter used to symbolize the quantity and its ordinal position in the alphabet.
	Letter as Representing Variable with Fixed but Arbitrarily Chosen Value	Recognizes a variable quantity as an unknown with a single, fixed value *that can be randomly chosen;* Sees literal symbol as representing "any number," but conceives of "any number" as *any fixed number* that is arbitrarily chosen.
	Letter as Representing Variable as Varying Unknown	Recognizes variable quantity as varying unknown and literal symbol as representing a varying, unknown quantity.

Reification	*Letter as Representing Variable as Mathematical Object*	Able to construct functional relationships that appropriately represents variables and constants; Able to act on variables, represented with literal symbols, as tools for representing generalized functional relationships.

Table 1. Trajectory in children's thinking about variable and variable notation.

Conclusion

In this study, we have tried to characterize a learning trajectory in young children's thinking about variable and variable notation. As with other research whose goal is to identify trajectories in children's thinking, our purpose has not been to quantify "how many" children exhibited thinking at a particular level. Instead, it has been to look closely at children's thinking as they encountered variable and variable notation through functional thinking tasks and map out levels of sophistication that emerged in their thinking in response to our instructional sequence. Even so, we do not want to lose sight of the extent to which some children in the study were able to understand variable and variable notation. It is noteworthy on its own that *any* of the six-year-old participants exhibited thinking characteristic of Level 6. In our view, this underscores points raised in other research regarding adolescents' difficulties with variable notation and whether or not these difficulties might be attributed more to how the notation conflicts with the many and varied experiences students have with literal symbols and less to one's age (McNeil et al., 2010). Based on the results of our study, we would further argue that adolescents' difficulties with variable and variable notation do not necessarily portend that younger children will have even greater difficulties with these concepts. In fact, the opposite might be true.

Moreover, we recognize that some children in our study—like adolescents in other research—took up symbolic notation early on as a way of representing an *object* they perceived to be unknown (e.g., a person). This misalignment between symbol and referent seemed rooted in a logical attempt by children to make sense of what the literal symbol might represent *in the absence of perceiving a variable*, leading to what has been characterized as object/quantity confusion and which has been interpreted as an inherent difficulty with a symbolic system (Lucariello & Tine, in press) and one's developmental readiness (Küchemann (1978, 1981). We suggest, however, that the root of this issue might not be in the child's lack of ability to use a symbolic system, but in the fact that he or she does not yet perceive the unknown quantity to be mathematized. Furthermore, we contend that the difficulty students—even adolescents—exhibit might be more closely attributed to the lack of experiences children are given to mathematize situations that involve unknown quantities than to developmental constraints that cast them as "not ready" to use symbols. We found that once children perceived a variable quantity, they were able to advance in their thinking about variable and variable notation in quite sophisticated ways. We wonder, therefore, what the effect would be if children had routine experiences in which they mathematized problems with unknown quantities, and whether these experiences might more naturally promote and support children's development of variable notation.

While we do not claim that students in this study developed a full and complete understanding of variable or variable notation, we do think our findings hold potential for the argument that young children *can* begin to think about these core algebraic areas. Furthermore, we suggest that providing long-term, sustained experiences with variable and variable notation from the start of formal schooling has the promise of ameliorating the deeply held misconceptions that adolescents exhibit.

We recognize, however, that some view the use of variable notation in lower elementary grades as controversial or even unnecessary. While we strongly support that children can use their own natural language to reason about algebraic situations, we equally support offering them other representations—such as variable notation—as a way to represent and reason with their ideas. We wonder how children would cope with natural language—also a symbolic and highly complex system with its own rules and syntax—if they were not exposed to squiggles to which we ascribe alphabetic meaning as "letters" long before they start formal schooling. We wonder how they might fare if they were not given experiences until middle grades to begin combining these letters in sequences to form words, then sentences, through an (arbitrary) system of rules and syntax that governed these actions and sometimes exhibited little logic (try explaining to a 6-year old why "might" is not spelled "mite" in the sentence "We might go."). We suggest that the development of natural language would be chaotic and littered with tales of students' misconceptions and difficulties. Yet, while developing one's natural language in this way would be unthinkable, it has been the norm for developing children's algebraic language. Our goal in identifying trajectories in children's thinking about variable and variable notation is to help develop a foundation for designing instructional experiences that can nurture children's thinking in this area and help shift perceptions that algebraic language is beyond the grasp of young children.

References

Baroody, A., Cibulskis, M., Lai, M., Li, X. (2004). Comments on the use of learning trajectories in curriculum development research. *Mathematical Thinking and Learning 6*(2), 227-260.

Battista, M. (2004). Applying cognition-based assessment to elementary school students' development of understanding of area and volume measurement. *Mathematical Thinking and Learning 6*(2), 185–204.

Blanton, M., Brizuela, B., Gardiner, A., Sawrey, K., Newman-Owens, A. (2014). A learning trajectory in six-year-olds' thinking about generalizing functional relationships. Manuscript submitted for publication.

Blanton, M., Levi, L., Crites, T., & Dougherty, B. (2011). *Developing Essential Understanding of Algebraic Thinking for Teaching Mathematics in Grades 3-5*. Essential Understanding Series. Reston, VA: National Council of Teachers of Mathematics.

Blanton, M., Stephens, A., Knuth, E., Gardiner, A., Isler, I., Kim, J. (2013). The development of children's algebraic thinking: The impact of a comprehensive early algebra intervention in third grade. Manuscript under review.

Booth, L. R. (1988). Children's difficulties in beginning algebra. In A. F. Coxford & A. P. Schulte (Eds.), The ideas of algebra, K–12 (pp. 20–32). Reston, VA: National Council of Teachers of Mathematics.

Braddon, K. L., Hall, N. J., & Taylor, D. (1993). Math through children's literature: Making the NCTM standards come alive. Englewood, CO: Teacher Ideas Press.

Clements, D. H. & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning, 6*(2), 81-89.

Clements, D., & Sarama, J. (2009). *Learning and teaching early math: The learning trajectories approach*. New York: Routledge.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9-13.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design Research: Theoretical and Methodological Issues. *The Journal of the Learning Sciences, 13*(1), 15-42.

Confrey, J., Maloney, A., & Nguyen, K. (2011) Learning over time: Learning trajectories in mathematics education. Charlotte, NC: Information Age.

Cooper, T., & Warren, E. (2011). Years 2 to 6 students' ability to generalize: Models, representations, and theory for teaching and learning. In J. Cai & E. Knuth (Eds.), *Early algebraization: A global dialogue from multiple perspectives* (pp. 187-214). Netherlands: Springer.

Glaser, B. G. (1998). *Doing grounded theory: Issues and discussions*. Mill Valley, CA: Sociology Press.

Glaser, B. G. (2002). Conceptualization: On theory and theorizing using grounded theory, *International Journal of Qualitative Methods, 1*(2), 23-38.

Kline, M. (1972). *Mathematical thought from ancient to modern times*. New York: Oxford University Press.

Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. C. (2005). Middle school students' understanding of core algebraic concepts: Equality and variable. International Reviews on Mathematical Education, 37, 1–9.

Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. C. (2011). Middle school students' understanding of core algebraic concepts: Equivalence and variable. In

J. Cai & E. Knuth (Eds.), *Early algebraization: A global dialogue from multiple perspectives. Advances in Mathematics Education Monograph Series* (pp. 259-276). New York: Springer.

Küchemann, D. E. (1981). Algebra. In K. Hart (Ed.), *Children's Understanding of Mathematics* (pp. 102-119). London, UK: Murray.

Lesh, R., & Yoon, C. (2004). Evolving communities of mind—In which development involves several interacting and simultaneously developing strands. *Mathematical Thinking and Learning, 6*(2), 205-226.

Lucariello, J., & Tine, M. (in press). Algebraic misconceptions: A teacher use test for diagnosing student misconceptions of the variable. In N. L. Stein & S. Raudenbush (Eds.), Developmental science goes to school. New York, NY: Taylor & Francis.

MacGregor, M., & Stacey, K. (1997). Students' understanding of algebraic notation: 11–15. Educational Studies in Mathematics, 33, 1–19.

McNeil, N., Weinberg, A., Hattikudar, S., Stephens, A.C., Asquith, P., Knuth, E., & Alibali, M. (2010). A is for apple: Mnemonic symbols hinder the interpretation of algebraic expressions, *Journal of Educational Psychology 102*(3), 625-634.

Moss, J., Beatty, R., Shillolo, G. & Barkin, S. (2008). What is your theory? What is your rule? Fourth graders build their understanding of patterns and functions on a collaborative database. In C. Greenes (Ed.), *Algebra and Algebraic Thinking in School Mathematics: The National Council of Teachers of Mathematics 70th Yearbook* (pp. 155-168). Reston, VA: National Council of Teachers of Mathematics.

Moss, J., & McNab, S. (2011). An Approach to Geometric and Numeric Patterning That Fosters Second Grade Students' Reasoning and Generalizing About Functions and Co-Variation.

In J. Cai & E. Knuth (Eds.), *Early Algebraization: A Global Dialogue from Multiple Perspectives. Advances in Mathematics Education Monograph Series* (pp. 277-301). Heidelberg, Germany: Springer.

Piaget, J. (1964). *Developing and learning: Conference on cognitive studies and curriculum development.* Ithaca, NY: Cornell University Press.

Resnick, L. B. (1982). Syntax and semantics in learning to subtract. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 136-155). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sarama, J., & Clements, D. H. (in press). Learning trajectories: Foundations for effective, research-based education. In J. Confrey, A. P. Maloney & K. Nguyen (Eds.), *Learning over time: Learning trajectories in mathematics education*. New York, NY: Information Age.

Schwartz, J. (1990). Getting students to function in and with algebra. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy*, (pp. 261-289). Washington, DC: Mathematics Associations of America.

Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26,* 114–145.

Stevens, S. Y., Shin, N., & Krajcik, J. S. (2009). *Towards a model for the development of an empirically tested learning progression.* Paper presented at the Learning Progressions in Science (LeaPS) Conference, Iowa City, IA.

Strauss, A.L., & Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Los Angeles, CA: Sage.

Vygotsky, L. (1978). Mind in society. (M. Cole, S. Scribner, V. John-Steiner, & E. Souberman, Trans.). Cambridge, MA: Harvard University. (Original work published in 1934.)

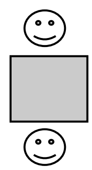
APPENDIX

Lesson 4: Attached Tables (adapted from Blanton et al., 2013)

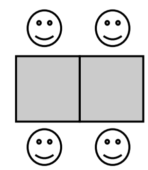
1. Brady's Birthday

Brady is celebrating his birthday at school. He wants to make sure he has a seat for all of his friends. He has square desks for his friends, but no one can sit on the ends of the desks.

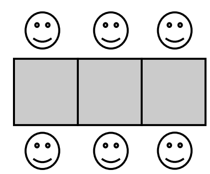
He can seat 2 friends at one desk in the following way:



If he joins another desk to the first one, he can seat 4 friends:



If he joins another desk to the second one, he can seat 6 friends:



Question: Can we find a relationship between the number of desks and the number of friends that can be seated?

A. Explore with Partner: How many friends could be seated at 10 desks? How did you figure this out?

B. Class Discussion:

- How many friends can be seated at 1 desk?
- How many friends can be seated at 2 desks? (continue for 3, 4 and 5 desks)

(Create a function table to organize the information. Discuss the meanings within this representation.)

number of desks	number of friends
1	2
2	4
3	6
4	8
5	10

C. Explore with Partner: Do you see any patterns or relationships here? How would you describe these?

D. Class Discussion:

Discuss the relationships students see.

- If Brady keeps adding desks, what will happen to the number of friends he can seat? How would you describe this?
- What happens to the number of friends every time we add one more desk?
- How would you complete the following:

  Every time the number of desk goes up by____, the number of friends goes up by _____.

- Can we add information to our chart even without having people come up to the board?
- How would I figure out the number of friends for 100 desks

3. Develop a Function Rule

Class Discussion:

Do you see any relationship between the number of desks and number of friends? How would you describe it?

*Discuss **how** students got each total number of friends and write an equation representing a relationship between each two corresponding values in the table.*

number of desks (D)	number of friends (F)	
1	2	1 + 1 = 2
2	4	2 + 2 = 4
3	6	3 + 3 = 6
4	8	4 + 4 = 8
5	10	5 + 5 = 10

Explore with Partner:

- If Brady has 20 desks, how many friends can he seat? How did you get your answer?
- If Brady wants to have 30 friends at his party, how many desks will he need? How did you get your answer?

[1] The research reported here was supported in part by the National Science Foundation under DRL 1154355. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

[2] By *variable notation*, we refer to literal symbols used to represent variable quantities. By *variable* or *variable quantity*, we refer to both varying and fixed unknown quantities (Blanton, Levi, Crites, & Dougherty, 2011).

[3] We use *literal symbol* interchangeably with *letter*.