Title
How Counting Represents Number: What Children Must Learn and When They Learn It

Journal
Cognition, in press

Authors
Sarnecka, Barbara W
Carey, Susan

Publication Date
2008

Peer reviewed

How Counting Represents Number: What Children Must Learn and When They Learn It

Barbara W. Sarnecka

University of California, Irvine

Susan Carey

Harvard University

Author Note

Barbara W. Sarnecka, Department of Cognitive Sciences, University of California, Irvine; Susan Carey, Department of Psychology, Harvard University; Michael D. Lee, Department of Cognitive Sciences, University of California, Irvine.

This research was supported in part by NSF REC-0337055 to Elizabeth Spelke and Susan Carey. We are grateful to Michael D. Lee for his help with the Bayesian analysis of the Last-Word data. We thank the children who participated in the study and their families, as well as the teachers and administrators of the following Irvine, California preschools: Jenny Hart Early Education Center, Temple Bat Yahm Preschool, Turtle Rock Preschool, UCI Children's Center, UCI Early Childhood Education Center, University Montessori, and Verano Preschool. We also thank research assistants John Cabiles, Alexandra Cerutti, Hillary McManama, Jyothi Ramakrishnan, Dat Thai and Gowa Wu for their work on data collection and coding; and lab managers Lacey Beckman, Emily Carrigan and Bailey House for their help with subject recruitment and scheduling. Thanks also to the members of the Harvard Lab for Developmental Studies and the UC-Irvine Cognitive Development Lab for their thoughtful comments at all stages of the project.

Correspondence concerning this article should be addressed to Barbara W. Sarnecka, Department of Cognitive Sciences, SSPA 3151, University of California, Irvine, California 92617-5100. E-mail: sarnecka@uci.edu.

Abstract

This study explored the conceptual basis for the cardinal principle of counting. The Give-N task was used to separate 73 2- to 4-year-olds into children who could give the right number of items for only a subset of the numerals in their count list ("subset-knowers") and children who could give the right number for all numerals tested ("high-numeral knowers"). Performance on two novel tasks supported the hypothesis that only the high-numeral-knowers understand how counting implements the successor function. Other tasks established that subset-knowers have good procedural competence with counting, and that a high percentage of subset-knowers know that the last word reached in a count is the appropriate answer to a "how many?" question. The results add to a body of literature detailing the many steps involved in working out how counting represents natural number.

How Counting Represents Number: What Children Learn and When

It seems uncontroversial to say that some children know how to count and others don't. But how to tell them apart? If knowing how to count just means reciting the numeral[1] list (i.e., "one, two, three . . . ") up to "five" or "ten," perhaps pointing to one object with each numeral, then many two-year-olds count very well (Baroody & Price, 1983; Briars & Siegler, 1984; Fuson, Richards, & Briars, 1982; Fuson, 1988; Gelman & Gallistel, 1978; Miller & Stigler, 1987; Schaeffer, Eggleston, & Scott, 1974). That kind of counting is good for marking time (e.g., close your eyes and count to ten . . .) or for playing with one's parents, but reciting the alphabet or playing patty-cake would do just as well. The thing that makes counting different from reciting the alphabet or playing patty-cake is that only counting tells you the number of things in a set.

Of course, counting only tells you this if you do it correctly, following the three 'how-to-count' principles identified by Gelman and Gallistel (1978). These are: The one-to-one principle, which says that "in enumerating (counting) a set, one and only one [numeral] must be assigned to each item in the set." (p. 90); the stable-order principle, which says that "[Numerals] used in counting must be used in the same order in any one count as in any other count." (p. 94); and the cardinal principle, which says that "the [numeral] applied to the final item in the set represents the number of items in the set."

---

[1] The words "one, two, three, . . ." etc. are commonly called "number words" in the psychological literature and "numerals" in linguistics. We prefer "numeral" because it means a symbol for an integer, whereas many other words (e.g., *few, several, many,* etc.) convey information about number, and so could be considered "number words."

(p. 80).

As Gelman and Gallistel pointed out, so long as the child's counting obeys these three principles, the numeral list ("one," "two," "three,". . . etc.) represents the cardinal numbers 1, 2, 3, . . . etc., as generated by the successor function. If numeral "n" represents cardinality *n*, then the next numeral on the list represents the cardinality "n + 1," which is the successor of n. The counting principles are what make counting equivalent to saying "one" (and one is) "two" (and one is) "three" . . .

In their 1978 book, Gelman and Gallistel argued that even two-year-olds honor these principles when counting, because the principles are intuitively understood and need not be learned. This view has come to be called the *principles-first* (or *principles-before-skills*) view.

Other studies, however, have failed to provide support for the principles-first hypothesis. For example, three-year-old children violate the one-one principle by skipping or double-counting items, or by using the same numeral twice in a count (Baroody & Price, 1983; Briars & Siegler, 1984; Frye, Braisby, Lowe, Maroudas, & Nicholls, 1989; Fuson, 1988; Miller, Smith, Zhu, & Zhang, 1995; Schaeffer et al., 1974; Wagner & Walters, 1982). Children also frequently violate the stable-order principle by producing different number-word sequences at different times (Baroody & Price, 1983; Frye et al., 1989; Fuson et al., 1982; Fuson, Secada, & Hall, 1983; Miller et al., 1995; Wagner & Walters, 1982). These findings have led many observers to conclude that the how-to-count principles, rather than being understood intuitively, are in fact learned gradually. This is known as the *principles-after* (or *skills-before-principles)* view.

Of course, as Greeno et al. (1984) point out, children might make these mistakes even if they do understand how counting represents number. Much more troubling for the principles-first hypothesis is evidence that young children do not understand the cardinal principle. That is, children do not seem to recognize that the last numeral used in counting tells the number of items in the set. One type of evidence comes from How-Many tasks. The version used by Schaeffer et al. (1974) is typical:

> "Each child was asked to count the chips in a line of *x* poker chips, where *x* varied between 1 and 7. After the child had counted the chips, the line was immediately covered with a piece of cardboard and the child was asked how many chips were hidden. Evidence that he knew . . . [the cardinal principle] was that he could respond by naming the last [numeral] he had just counted." (p. 360)

Studies using a How-Many task have found that many children respond incorrectly to the "how many" question even after they have counted the array correctly. When asked “how many,” children often try to count the set again, or (if prevented from recounting) either make no response or guess some numeral other than the last numeral in the count sequence (Frye et al., 1989; Fuson, 1992; Markman, 1979; Rittle-Johnson & Siegler, 1998; Schaeffer et al., 1974; Wynn, 1990; Wynn, 1992). Moreover, some investigators have concluded that the How-Many task *over*estimates children’s knowledge, because some children actually do repeat the last numeral used in counting without (apparently) understanding that it refers to the cardinal value of the set—this superficial understanding has been called a 'Last-Word Rule' to distinguish it from the cardinal principle (Frye et al., 1989; Fuson, 1988).

Conversely, supporters of the principles-first position argue that How-Many tasks generally demonstrate that children *do* understand the cardinal principle. They say that, if anything, this type of task *under*estimates children's knowledge, because it is pragmatically strange to ask "how many" immediately after counting (Gelman, 1993; Greeno et al., 1984). To demonstrate this, Gelman (1993) did a How-Many task with college undergraduates: "When we asked undergraduates a how-many question about 18 blocks, all of them counted but only one bothered to repeat the last count word said. Repeats of the question elicited puzzlement, some recounting, and so forth . . . " (p. 80).

In short, How-Many tasks seem to raise as many questions as they answer: What must a child know in order to succeed on a How-Many task? And how does that knowledge relate to other things the child knows about counting and number, such as the meanings of the numerals, or the how-to-count principles?

The Give-N task provides a different kind of evidence that young children do not understand the cardinal principle. (This task is also called "Give-a-Number" by Wynn, 1990, 1992 and others; "Give-me-X" by Frye et al., 1989; and "Make a set of N" by Fuson, 1988). In this task, the child is asked to create a set with a particular number of items. For example, the experimenter might ask the child to "Give two lemons" to a puppet. (In a slightly different version of the task, Schaeffer et al., 1974 asked children to tap a drum *n* times.) Studies using these tasks have found that children are often unable to create sets for numerals that are well within their counting range. (E.g., many children who can count to "five" cannot create sets of five objects.)

Levels of Performance on the Give-N Task

In addition to providing evidence that young children do not understand the cardinal principle, Give-N studies have yielded a detailed picture of the learning pattern for numeral meanings. It turns out that a child's performance on the Give-N task goes through a series of predictable levels, found in a longitudinal study by Wynn (1992) and in several cross-sectional studies (Condry & Spelke, in press; Le Corre, Van de Walle, G., Brannon, & Carey, 2006; Le Corre & Carey, in press; Sarnecka & Gelman, 2004; Sarnecka, Kamenskaya, Yamana, Ogura, & Yudovina, in press; Schaeffer et al., 1974; Wynn, 1990). These performance levels are found in child speakers of Japanese (Sarnecka et al., in press) Mandarin Chinese (Le Corre, Li, & Jia, 2003; Li, Le Corre, Shui, Jia, & Carey, 2003) and Russian (Sarnecka et al., in press) as well as English.

The developmental pattern is as follows. At the earliest level, the child makes no distinctions among the meanings of different numerals. On the Give-N task, she may always give one object to the puppet or she may always give a handful, but the number she gives is unrelated to the numeral requested. A child at this level can be called a ”no-numeral-knower,” for she has not yet assigned numerical meaning to any of the numerals in her count list (i.e., the list of numerals that she has memorized).

At the next level (which most English-speaking children reach by age 2-1/2 to 3 years) the child knows only that “one” means one. On the Give-N task, she gives one object when asked for “one,” and she gives two or more objects when the request is made with any other numeral. This is the “one”-knower level.

Some months later, the child becomes a “two”-knower, for she learns that “two” means two. At that point, she gives one object when asked for “one,” and two objects when asked for “two,” but she does not distinguish among the numerals “three,” "four," "five," etc. For any of those numerals, she simply grabs some objects and hands them over. This is followed by a “three”-knower level, and some studies also report a “four”-knower level. Collectively, children at these levels have been termed “subset-knowers”

(Le Corre et al., 2006; Le Corre & Carey, in press) because although most of them have memorized the numeral list up to “ten” or higher, they can generate sets to match only a subset of those numerals.

After the child has spent some time (often more than a year) as a subset-knower, she moves on to the next level. Now she is able to generate the right set for a numeral (on the Give-N task) or the right the numeral for a set (on the ‘What's-On-This-Card’ task, Gelman, 1993; Le Corre et al., 2006) for the numerals "five" and above. But whereas she progressed through the subset-knower levels gradually (learning "one," then "two," then "three" . . .) she seems to acquire the meanings of the higher numerals ("five" through however high she can count) all at once. We will call children at this level *high-numeral-knowers*.

Within-child consistency on a wide variety of tasks suggests that high-numeral-knowers differ qualitatively from subset-knowers. For example, a “two”-knower is, by definition, unable to give a puppet three or more items on the Give-N task. But a "two"-knower is *also* (a) unable to fix a set when told for example, "Can you count and make sure you gave the puppet three toys? . . . But the puppet wanted *three*-- Can you fix it so there are three?" (Le Corre et al., 2006) (b) unsure whether a puppet who has counted out seven items has produced a set of “seven” (Le Corre et al., 2006); (c) unable to point to the card with “three” apples, given a choice between a card with 3 and a card with 4 (Wynn, 1992); and (d) unable to produce the numeral "three" when presented with a set of 3 items on a card (Le Corre et al., 2006). High-numeral-knowers succeed across the board on these tasks.

Qualitative differences in the counting behavior of subset-knowers and high-numeral-knowers strongly suggest that what ultimately separates these groups is knowledge of the cardinal principle. Most conspicuously, subset-knowers do not use counting to generate sets on the Give-N task, whereas high-numeral-knowers do (an observation that led Wynn, 1990, 1992 to call subset-knowers "grabbers," and high-numeral-knowers "counters").

What Do High-Numeral-Knowers Know?

According to the principles-after view, the difference between high-numeral-knowers and subset-knowers is that high-numeral-knowers have induced how counting implements the successor function (i.e., the function that generates each integer by adding 1 to the integer before it) whereas subset-knowers have not. Carey (2004), Hurford (1987) and Klahr (Klahr & Wallace, 1976; 1984) develop closely related proposals for a learning mechanism that could achieve this feat. But this hypothesis about the fundamental difference between high-numeral-knowers and subset-knowers has never been directly tested, and it is this gap in the literature the present study is designed to close.

If children know how counting represents the successor function, they should understand two things. (1) The *direction* of numerical change: The numeral that denotes cardinality N+1 will be somewhere after the numeral denoting cardinality N in the numeral list. (2) The *unit* of numerical change: The numeral for cardinality N+1 must be the very next numeral in the list after the numeral for cardinality N.

The Present Study

We devised two simple measures (the Direction task and the Unit task) to tap children's understanding of how the direction and unit of numerical change are represented by moving forward or backward along the numeral list. If the real difference between high-numeral-knowers and subset-knowers is that high-numeral-knowers have induced how counting implements the successor function, then high-numeral-knowers should succeed at these two tasks and subset-knowers should fail. If subset-knowers and high numeral-knowers differ in some other way, then the two groups should not be distinguishable by their performance on the Direction and Unit tasks. These tasks are especially suitable because they involve addition and subtraction from sets, which according to Gelman and colleagues is the best way to reveal children's conceptual competence with respect to number representation (Cordes & Gelman, 2005; Zur & Gelman, 2004).

We also devised a How-Many task that circumvents the pragmatic oddness of asking the child to say how many items there are in a set they have just counted. The experimenter counted a set the child could not see, and then asked the child to say how many there were in the set. This task allowed us to assess when children learn a last-word rule (i.e., a rule saying that the answer to a "how many?" question is the last word in a count) and how mastery of this rule relates to knowledge of the cardinal principle as assessed by Give-N.

The strategy of the present study thus involved testing each child on six different tasks. First, the Give-N task was used to divide children into no-numeral-knowers, subset-knowers ("one," "two," "three" and "four"-knowers) and high-numeral knowers who can give the puppet up to "six" items when requested. Next, two tests of counting fluency (the Sequence and Correspondence tasks) were included to provide a baseline measure of the child's mastery of the numeral sequence and standard counting procedure.

Then, the Last-Word task was used to probe for a last-word rule as described by Fuson (1988) and Frye et al. (1989). Finally, two new tasks (the Direction and Unit tasks) tested whether children knew that (a) adding an element to a set requires going forward in the numeral list to represent the number in the resulting set, while subtracting requires going backward (the Direction task) and (b) it is the next numeral that represents the set if 1 item has been added and the numeral after that if 2 have been added (the Unit task).

Method

Participants

Participants included 73 children (28 boys, 45 girls), ranging in age from 2 years, 10 months to 4 years, 3 months (mean age 3-6). All children were monolingual and native speakers of English. Approximately half the children (38 out of 73) were recruited by mail and phone using public birth records in the greater Boston, Massachusetts area. These children were tested at a university child development laboratory. Parents who brought their children in for testing received reimbursement for their travel expenses and a token gift for their child. The other 35 children were recruited and tested at university-affiliated or private child-care centers in Irvine, California.

Parents received a token gift when they signed their child up for participation; preschools received gift certificates to a children's bookstore and cognitive development seminars for their staff. Families were not asked about their ethnicity, household income, or education, but participants were presumably representative of the middle-class, predominantly European-American and Asian-American communities in which they lived. The two samples (Massachusetts and California) did not differ significantly in age or in proportion of girls to boys.

Children were tested in one or two sessions, depending on their willingness to keep playing. In order for a child's data to be included, she had to complete at least three of the six tasks in the study. Three additional children (two girls, ages 32 months and 40 months; and one boy, age 38 months) were tested but asked to stop playing before they completed three tasks. These children's data were excluded.

Tasks

Give-N task (Frye et al.,1989; Wynn, 1990, 1992). The purpose of this task was to determine which numerals the child knew the exact meanings of. How a child performed on this task determined her 'knower-level' (i.e., "one"-knower, "two"-knower, high-numeral-knower, etc.). Materials for this task included a green dinosaur puppet (approx. 24 cm tall and 24 cm in circumference), a blue plastic plate (approx. 11 cm in diameter), and 15 small plastic lemons (approx. 2 x 3 cm each). To begin the task, the experimenter placed the puppet, plate, and lemons on the table and said, "In this game, you will give things to the dinosaur, like this." (The experimenter mimes placing something on the plate, then slides the plate over to the puppet.) Requests were of the form "Can you give *one* lemon to the dinosaur?" After the child responded to each request, the experimenter asked the follow-up question, of the form "Is that *one*?" If the child said no, the original request was restated (e.g., "Can you give the dinosaur *one* lemon?"), followed again by the follow-up question (e.g., "Is that *one*?"). This continued until the child affirmed that the dinosaur had the requested number of objects.

All children were first asked for one lemon, then three lemons. Further requests depended on the child's earlier responses. When a child responded correctly to a request for N, the next request was for N+1. When she responded incorrectly to a request for N, the next request was for N-1. The requests continued until the child had at least two successes at a given N (unless the child had no successes, in which case she was classified as a no-numeral-knower) and at least two failures at N+1 (unless the child had no failures, in which case she was classified as a high-numeral-knower).

The highest numerals requested were "five" and "six." It was important to include the numeral "five" because earlier studies have reported that children who can generate sets of five in the Give-N task are "counters" rather than "grabbers" (Wynn, 1990) or "CP-knowers" rather than "subset-knowers" (Le Corre & Carey, in press; Le Corre et al., 2006). But early testing showed that a handful of lemons for many of the children happened to be five, meaning that they often generated sets of five just by chance. Also, we reasoned that "four"-knowers might be likely to generate sets of five just by adding lemons to the plate until the set size was bigger than they could name. For these reasons, we requested "five" and "six" in alternation (i.e., all children who received two high-numeral requests got one request for "five" and another for "six").

A child was credited with knowing a given numeral if she had at least twice as many successes as failures for that numeral. Failures included either giving the wrong number of items for a particular numeral *N*, or giving N items when some other numeral was requested. The highest numeral each child succeeded at determined her knower-level. (For example, children who succeeded at "one" and "two," but failed at "three" were called "two"-knowers.) Children who had at least twice as many successes as failures for high-numeral trials (i.e., trials of "five" and "six") were called high-numeral-knowers.

Sequence task. This task measured the child's knowledge of the numeral up to “ten.” To begin the task, the experimenter said, "Let's count. Can you count to ten?" If the child did not immediately start counting, the experimenter said, "Let’s count together. One, two, three, four, five, six, seven, eight, nine, ten. OK, now you count." Each child's score reflects the highest numeral she reached without errors. For example, a child who counted "one, two, three, four, five, six, seven, nine, ten" counted correctly to 'seven,’ and so received that score. Children were allowed to start over if they asked to, or if they did so spontaneously, but they were not told to start over by experimenters. For children who counted more than once, only their best count was used.

Correspondence task. This task measured the child’s skill at tagging objects in one-to-one correspondence with the numeral list. Materials included two cork boards (36 cm x 12.5 cm) and a set of large, brightly colored push pins (2.5 cm across). One cork board contained an array of five push pins, the other contained ten; the pins were evenly spaced and arranged in a straight line. To begin the task, the experimenter presented the array and said "Now show me how you count these." The array of five was always presented first, followed by the array of ten. Each child's score reflects the highest number of items she was able to count without skipping or double-counting. No errors were allowed on the array of five; a maximum of one error was allowed on the array of ten (i.e., a child who made only one skip or double-count error on the array of ten still received a score of ten). Children were allowed to start over if they asked to, or if they did so spontaneously, but they were not told to start over. For children who counted an array more than once, only their best count was used.

Last-Word task. This task probed children's responses to a how-many question following a standard count by an adult. Materials for this task included three picture cards (approx. 23 cm x 7 cm) each depicting a row of items (three peppers, each 7 x 7.5 cm; five onions, each 5 x 5 cm; and seven tomatoes, each 2.5 x 2.5 cm). The task was introduced as follows: The experimenter held the picture of three peppers facing away from the child, so that the child could not see it. The experimenter said, "I have a picture of some peppers here, and I'm going to count them. You try to guess how many peppers are in the picture by listening to my counting. Ready? One, two, *three*. OK, how many peppers?" The first trial was a training trial. The experimenter showed the card to the child afterward and commented on the child's answer (e.g., "That's right! It *was* three peppers." or "Oops, it was actually *three* peppers. Good try, though."). The point of this training trial was to demonstrate how the game worked, and to show that the experimenter had counted in a standard way (i.e., the experimenter was not 'tricking' the child by counting wrong).

Next, the test trials (using the pictures of five and seven items) were presented in counterbalanced order. On these trials, the child received mildly positive feedback (e.g., “okay”) regardless of her answer, and was not allowed to see the cards after answering.

Direction task. The purpose of this task was to find out whether the child understood that moving forward in the numeral list represents adding items to a set, whereas moving backward in the list represents subtracting items. Materials included two clear plastic plates (approx. 23 cm in diameter) and four small plastic tubs. Each tub contained 12 small objects, six of one color and six of another color (light blue and magenta hair bands, green and orange jacks, red and purple bears, or yellow and dark blue dragonflies). Order of item presentation was randomized by allowing the child to choose the set of items for each trial. The experimenter began each trial by placing either five or six items of the same color on each plate, saying for example, "OK, I'm putting FIVE bears on here, and FIVE bears on here . . . so this plate has five, and this plate has five." Then the experimenter moved one item from one plate to the other, saying, "And now I'll move one." In this example, one of the plates would now contain four red bears, and the other would contain five purple bears and one red bear. Next the experimenter would say "OK, now there's a plate with FOUR, and a plate with SIX. And I'm going to ask you a question about the plate with SIX. Are you ready? Which plate has SIX?" Each child received four trials: Two trials started with five items per plate, one trial asked about "four," the other about six"; the other two trials started with six items per plate, one asked about "five," the other about "seven." Each trial was scored correct or incorrect.

Unit task. The purpose of this task was to find out whether the child understood that the unit of numerical increase represented by moving from one numeral to the next on the list is exactly one item. Specifically, this task tests whether children know that moving forward one word in the list means adding one item to the set, whereas moving forward two words in the list means adding two items to the set. Materials for this task included a wooden box (17.5 x 12.5 x 5 cm) and six small plastic tubs. Each tub contained seven identical toys (frogs, bananas, worms, sea horses, fish, or rabbits). Order of item presentation was randomized by allowing the child to choose the tub for each trial.

The experimenter began the trial by placing a number of items in the box, saying for example, "OK, I'm putting FOUR frogs in here." Then the experimenter closed the box and asked the memory-check question "How many frogs?" If the child did not answer the memory-check question correctly (e.g., “four”), the experimenter said, "Oops, let's try again" and repeated the beginning of the trial. After the child answered the memory-check question correctly, the experimenter said "Right! Now watch . . . " and added either one or two more items. Then the experimenter asked the test question, of the form, *Now is it N+1, or N+2?* (e.g., "Now is it FIVE, or SIX?") Each child received two warm-up trials (1+1 item and 1+2 items) followed by four test trials (4+1, 4+2, 5+1, and 5+2) in counterbalanced order. For the trials beginning with one item, the test question was "Now is it TWO, or THREE?" For the trials beginning with four items, the question was "Now is it FIVE, or SIX?" For the trials beginning with five items, the question was "Now is it SIX, or SEVEN?" No feedback was given after any of the trials, although children could see the contents of the box when the experimenter opened it to return the items to their tub. Each trial was scored correct or incorrect.

Order of tasks

Order of tasks was randomized in the following way. The materials for each task were placed inside a large, opaque drawstring bag. (The Sequence, Correspondence, and Last-Word tasks were grouped together in one bag.) Each bag was a different color (red, blue, green or yellow). At the beginning of the session, the child was asked, “What game should we play first?” and was given a choice of the four bags. The child chose a bag (without looking inside it), and the experimenter did the task in that bag. When the task was done, the child was allowed to choose from the three remaining bags, and so forth. For 42 of the children, all the tasks were presented in this way. The other 31 children completed the Give-N task during one session as part of a larger project, and completed the other five tasks (Sequence, Correspondence, Last-Word, Direction, and Unit) during a second session, no more than one week later (mean 4.6 days later.)

Results

Merging of Massachusetts and California samples. Initial analyses revealed no significant differences between the Massachusetts and California samples on any measure, so data were merged for the analyses reported here.

Give-N task. This task measured children's knowledge of the exact meanings of the numerals "one" through "six"; children were sorted into knower-levels on the basis of their performance on this task. Figure 1 shows the age distribution of each knower-level.

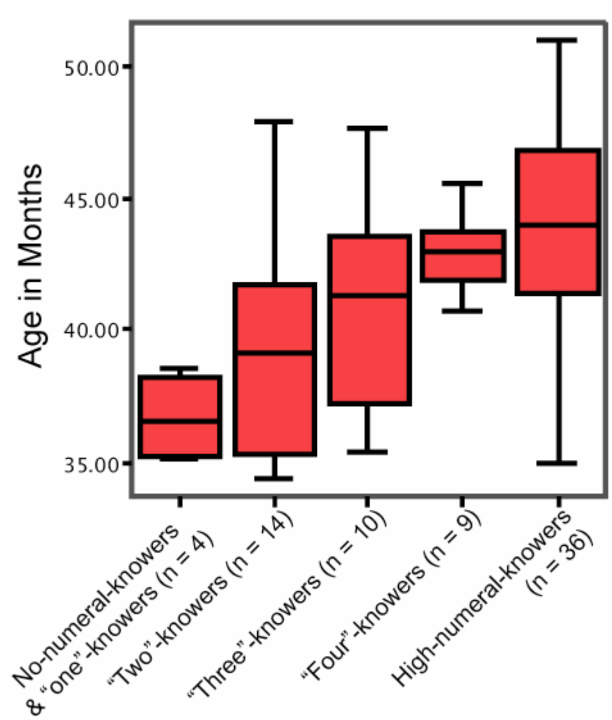

Figure 1. Ages of subjects at each knower-level. (Knower-levels determined by Give-N task.) Boxes enclose the middle 50% of values; horizontal line in box indicates the mean for each group; extending lines indicate the top 25% and bottom 25% of values.

Of the 73 children tested, 36 (49.3%) were high-numeral-knowers. These children ranged in age from 2-11 to 4-3 (mean age 3-8). The remaining 37 children (50.7%) were subset-knowers, so designated because they knew numerical meanings for only a subset of the numerals on their count list. Subset-knowers ranged in age from 2-10 to 4-0 (mean

3-5). Breaking it down by knower-level, there were two no-numeral-knowers (ages 2-11 and 3-3) and two “one”-knowers (ages 2-11 and 3-2); these groups were merged for analysis of performance on all other tasks (n = 4, mean age 3-1). There were also 14 “two”-knowers (ages 2-10 to 4-0, mean 3-3); ten “three”-knowers (ages 2-11 to 4-0, mean 3-5); and nine “four”-knowers (ages 3-5 to 3-10, mean 3-7).

Although we did not explicitly tell children to count, we did record whether or not children spontaneously counted out loud, either when giving lemons to the dinosaur or when answering the follow-up question, *Is that N*? Overall, high-numeral knowers were much more likely to count than subset-knowers, with 24 high-numeral-knowers (67%) counting aloud on at least one trial, whereas only 6 subset-knowers (16%) ever counted.

The distribution of counting across trial types is also informative. Subset-knowers were equally likely to count on low-numeral trials (i.e., trials asking for numerals "one," "two," "three," or "four") and on high-numeral trials (i.e., trials asking for "five" or "six”). Specifically, two of the subset-knowers who ever counted did so only on a single low-numeral trial; two counted only on one or two high-numeral trials, and two counted on both low- and high-numeral trials. Interestingly, three of the four subset-knowers who ever counted on one or more high-numeral trials were "four"-knowers. (The other was a "three"-knower.) Although these four children tried to use counting to construct sets of five and six items, they were not able to use it successfully. (None of these children succeeded on more than one high-numeral trial, and all of them failed at least two high-numeral trials – which is why they were classified as subset-knowers.)

The high-numeral-knowers, on the other hand, either used counting on the high-numeral trials only, or used it on both high- and low-numeral trials. (Unlike the subset-knowers, there was no high-numeral-knower who used counting only on low-numeral trials.) In fact, high-numeral-knowers were over six times more likely than subset-knowers to use counting on high-numeral trials—counting aloud on 67% of high-numeral trials, as compared to the subset knowers' 11%. (Setting "four"-knowers aside, only 1 out of 28 no-numeral-knowers-through-"three"-knowers ever counted, making high-numeral-knowers over 22 times more likely to count out large sets on the Give-N task than subset-knowers.)

Thus, our data find the same qualitative differences between high-numeral-knowers and subset-knowers as have been reported in earlier studies (Le Corre & Carey, in press; Le Corre et al., 2006; Sarnecka & Gelman, 2004; Wynn, 1990, 1992). The unit and direction tasks (see below) address whether the difference between high-numeral -knowers and subset-knowers is in fact that only high-numeral-knowers understand how counting implements the successor function.

Sequence and Correspondence tasks. These tasks measured participants’ skill at producing the numeral list (Sequence task) and at pointing to and counting arrays of objects (Correspondence task). On both tasks, children in all groups performed at or near ceiling, meaning that they recited the numeral list up to ten, and also pointed to each object in an array once and only once.

The mean score of all children on the Sequence task was 9.94 (range 8-10). The mean score and range of scores for each group (in order, beginning with the no-numeral-knowers/”one”-knowers) were 10.00 (range 10-10); 9.75 (range 8-10); 9.90 (range 9-10); 10.00 (range 10-10); and 10.00 (range 10-10), respectively.

The mean score of all children on the Correspondence task was 9.32 (range 1-10).

The mean score and range of scores for each group (in order, beginning with the no-numeral-knowers/"one"-knowers) were 8.75 (range 5-10); 9.07 (range 4-10); 8.80 (range 3-10); 9.44 (range 5-10); and 9.61 (range 1-10), respectively. The median and modal scores for each group on both tasks were 10.0.

All the children except two produced the numeral list up to 10 on at least one counting task. The exceptions were one "two"-knower (age 2-10) who produced the sequence up to 8, and one "three"-knower (age 4-0) who produced the sequence up to 9. Thus, every child was familiar with the portion of numeral sequence (i.e., "four-five-six-seven") relevant to the other tasks in this study. This baseline measure should be kept in mind when considering the rest of the analyses.

Last-Word task. The Last-Word task measured how often children answered a "how many" question with the last numeral of the experimenter's count. This was an open-ended task—a child could answer with the correct numeral, with a different numeral, or with no numeral at all. Since children knew at least ten numerals, if they were providing a numeral response but otherwise answering at random, chance would conservatively be 10%.

On the training trial (where the experimenter counted to three) 65 children (76%) correctly answered "three"; seven children (21%) produced some other numeral (answers ranged from "one" to "ten"), and one child (a "one"-knower) began counting out loud, starting at "one" and continuing until the experimenter stopped him at "twelve."

After feedback was given on the training trial, each child received two test trials, (where the experimenter counted to five and seven) in counterbalanced order. Overall, children answered correctly 83% of the time. On the other 17% of trials, children either responded with an incorrect numeral (13% of total trials) or counted aloud (4% of total trials). Children who counted always began at "one" and continued on past the target numeral—the exception was one trial where the experimenter had counted to "five" and the child also counted to "five" (i.e., parroted back what experimenter had said, word for word). This child counted to "five" on the trial of seven as well.

Breaking down performance by knower level: High-numeral-knowers almost always answered correctly (96% of trials). This was significantly higher than the subset-knowers' overall success rate of 68%, $t(67) = 3.53$, $p = .001$; see Table 1. However, all subset-knowers did not perform alike. On the contrary, the data in Table 1 show that the most dramatic difference was between the “one”-knowers and the “two”-knowers (25% correct and 64% correct, respectively). From “two”-knowers on, the success rate is always over 60%—the children were clearly not producing numerals at random.

Table 1

Results of Last-Word task

Knower-Level	n	Both trials correct	1 correct, 1 incorrect	Both trials incorrect
Pre- & One-Knowers	4	0	2	2
Two-Knowers	10	6	0	4
Three-Knowers	10	8	2	0
Four-Knowers	7	5	1	1
High-Number-Knowers	36	34	1	1

As is also apparent from Table 1, most children either got both trials right (53 children) or got them both wrong (8 children). It was relatively rare for a child to get one trial correct and the other incorrect. Of course, the high-numeral-knowers usually got both trials correct, but even among the subset-knowers, only 5 children got 1 trial right and 1 trial wrong, whereas 19 children got both trials right and 7 children got both trials wrong. This error pattern makes sense if there is a rule (i.e., a last-word rule) that has been learned by some children and not others. It is not the pattern one would expect to see if all participants understand the relevant concepts but sometimes do not apply them because of procedural error. A constant, low level of error would make it much more likely for any given child to get *one* trial incorrect than to get *both* trials incorrect.

To formally test these intuitions, we asked the following questions:

(1) Does the rate of correct answers increase with knower-level or stay the same?

(2) Is the rate of correct answering for each child essentially dichotomous (i.e., their probability of giving a correct answer is either very high or very low, so that they either get none right or both right)?

We developed four models that expressed all four combinations of these two possibilities, and estimated the Bayes Factors between these four models (to decide which model best fit the data) using the computational method known as 'reverse jump Markov chain Monte Carlo'. Results are given in Table 2. It is clear that the Dichotomous-Increasing model is by far the best one, with all of the Bayes Factors exceeding the suggested scientific standards for 'very strong' evidence (see Kass & Raftery, 1995, p. 777). That is, this analysis suggests that children either knew or didn't know the last-word rule, and that knowledge of the rule increased with knower-level.

Table 2

Bayes Factors between the four models, relative to the most likely Dichotomous-Increasing model.

Model	Bayes Factor	Log Bayes Factor
Dichotomous-Increasing	1	0
Dichotomous-Same	9×10^8	20.63
Continuous-Increasing	1,966	7.58
Continuous-Same	3×10^7	17.21

Note. Bayes Factors can be though of as betting ratios – e.g., the Continuous-Increasing model is 1,966 times less likely to be generated by the current data than the most likely Dichotomous-Increasing model. The Log Bayes Factors are included because they may be more familiar to some readers.

Thus, we see that many young children have formulated the generalization that the last word reached in a count is the appropriate answer to the question “how many.” Furthermore, they apply this rule long before they demonstrate any understanding of cardinality on tasks that don’t use the phrase “how many?” such as Give-N. In the present study, most subset-knowers succeeded on the Last-Word task, but on the Give-N task were able to generate sets for only some and not all of the numerals in their count list.

Direction task. This task measured children's understanding that counting *forward* in the numeral list represents *adding* items, whereas counting *backward* represents *subtracting* items. A one-way ANOVA actually found no significant main effect of knower-level on Direction scores, F(4) = 1.54, $p = .20$. However, a series of planned comparisons based on the original hypotheses showed that no-numeral-knowers/“one”-knowers, “two”-knowers and “three”-knowers all performed at chance, whereas “four”-knowers and high-numeral-knowers performed significantly above chance, $t(7) = 3.42$, $p = .01$ and $t(29) = 2.29$, $p = .03$, respectively. "Four"-knowers and high-numeral-knowers still made many errors, indicating that the task is a difficult one. (See Figure 2.)

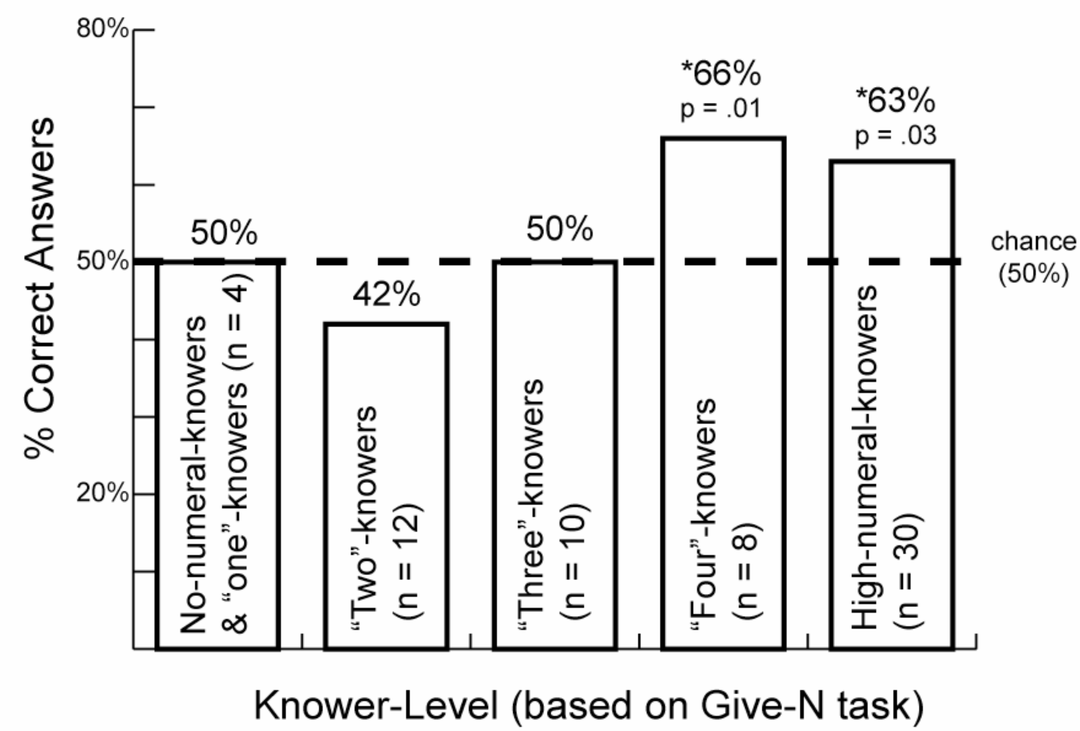

Figure 2. Direction task. Only “four”-knowers and high-numeral-knowers showed some understanding that moving forward in the numeral list corresponds to adding items, whereas going backward corresponds to subtracting items.

These results indicate that part of what separates high-numeral-knowers from subset-knowers is an understanding of the mapping between (a) the direction of movement along the numeral list and (b) the direction of change in the numerosity of a set. However, this can't be the whole story, because “four”-knowers succeed at the Direction task, but still don't use counting to solve the Give-N task (that's why they aren't high-numeral-knowers). Applying the cardinal principle must require some other piece of knowledge-- something that only the high-numeral-knowers know.

Unit task. This task measured children's understanding that going forward *one* word in the number list means adding *one* item, whereas going forward *two* words means adding *two* items.

In the first two (warm-up) trials, children were asked to judge whether a box that started with one item, and gained one or two more, had “two” items or “three.” No-numeral-knowers and ”one”-knowers performed at chance; every other group (i.e., “two”-knowers, “three”-knowers, “four”-knowers, and high-numeral-knowers) performed significantly above chance, *p*s < .05. These results from the warm-up trials indicate first, that children were able to understand the directions and do the task, and second that children at and above “two”-knower level have concepts of “one” and “two” that support the inferences required by this task. (As indeed we would expect them to.)

In the other four trials (“four” plus one, “four” plus two, “five” plus one and “five” plus two), a one-way ANOVA found a significant main effect of knower-level on Unit scores, $F(4) = 3.91$, $p < .01$. Planned comparisons based on the original hypotheses showed that no-numeral-knowers/”one”-knowers, “two”-knowers, “three”-knowers and “four”-knowers all performed at chance, whereas high-numeral-knowers performed significantly above chance, $t(28) = 4.61$, $p < .001$. This task, like the Direction task, was a difficult one, and even the high-numeral-knowers' performance was far from perfect. (See Figure 3).

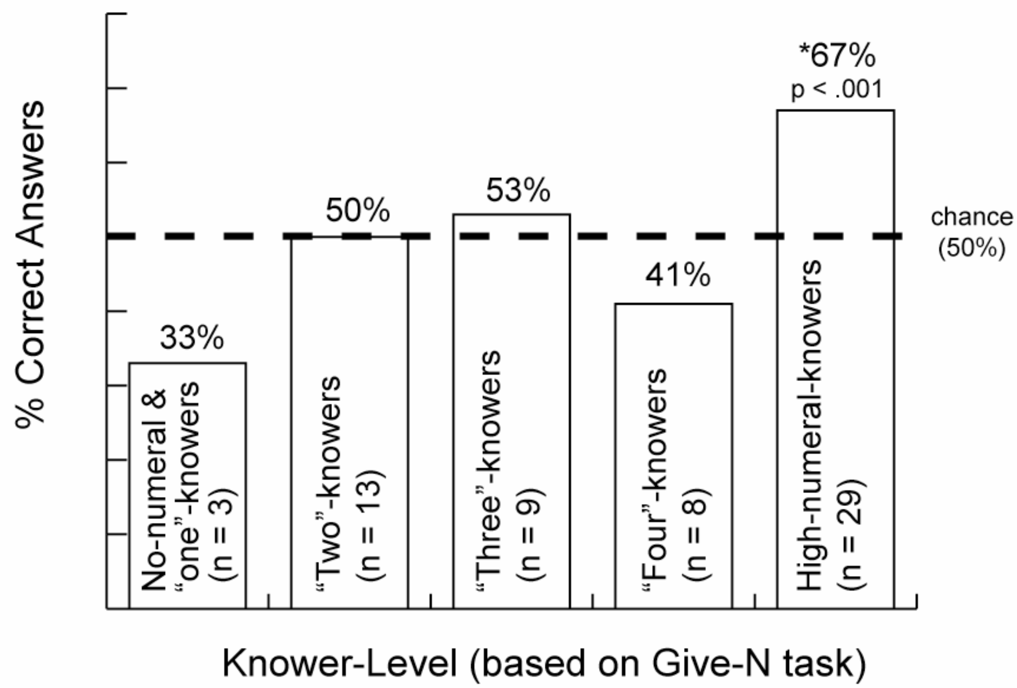

Figure 3. Unit task. Only high-numeral-knowers had any sense that going forward *one* word in the numeral list corresponds to adding *one* item, whereas going forward *two* words corresponds to adding *two* items..

These results suggest that what finally separates high-numeral-knowers from subset-knowers is an understanding of the *unit* of mapping between numerals and numerosities. Only high-numeral-knowers understand that going forward *one* word means adding *one* item; going forward *two* words means adding *two* items. In other words, the only children who know how (and when) to use counting in number-related problems like the Give-N task are those who understand (a) the mapping between *direction* of movement through the count list and *direction* of changes in numerosity, and

(b) the *unit* of change in numerosity signaled by movement from one numeral to the next on the list.

General Discussion

The jumping-off point for the present study was the (now well-established) observation that if you ask a two- to four-year-old child to give you various numbers of items (e.g., give me *two* blocks / give me *one* book / give me *four* crayons, etc.), that child's responses will fit the pattern for one of six knower-levels: No-numeral-knowers, "one"-, "two"-, "three"-, or "four"-knowers (i.e., subset-knowers), or high-numeral-knowers. Performance at different knower-levels varies along two parameters: (1) the set sizes children can generate upon request, and (2) whether or not they use counting to do it.

Why is it that some children (subset-knowers) can generate sets for only a few of the numerals in their count list (or none of them, in the case of no-numeral-knowers), while others generate sets for all the numerals we tested? And why is it that only the latter group uses counting (which seems the obvious way to solve the problem)? These are not original questions; they are the obvious questions to ask after observing the different levels of performance described here. But the present study brings new data to bear on these questions—replicating relevant findings from the literature and also presenting new findings from novel tasks designed specifically to answer the question: *What is it, exactly, that high-numeral-knowers know?*

First, we sorted children into knower-levels, based on their ability to create sets of one to six items upon request. Our data confirmed previously published descriptions of knower-levels, both in the set sizes children could generate and in the fact that only high-numeral-knowers regularly used counting to solve the problem.

Second, we tested children on two baseline measures of counting skill: Counting out loud up to ten (the Sequence task) and counting arrays of 5 and 10 objects (the Correspondence task). Of the 73 children tested, 71 counted to "ten" on at least one trial; every child in the study counted to "eight" or higher, indicating that they had mastered the numeral sequence in the range "four" through "seven" (the part of the sequence needed for the other tasks in the study).

Third, we devised a task to probe for a superficial last-word rule. In our task, the experimenter counted a set the child could not see and then asked the child to guess how many were in the set. Since random responding (choosing a numeral from one's count list at random) would lead to a probability of .1 or less, getting two responses correct cannot be due to chance. Virtually all of our high-numeral-knowers got both trials correct, but so did three-quarters of the "two," "three"- and "four"-knowers, and a quarter of the no-numeral-knowers. Clearly, knowledge of this rule is not equivalent to understanding the cardinal principle. These findings confirm earlier reports by Fuson (1988) and by Frye, et al. (1989) that before children understanding how counting represents number, they learn a superficial last-word rule: the answer to a "how many" question is the last word reached in a count.

Subset-knowers' relative success on our Last-Word task contrasts sharply with their complete failure on Le Corre et al.'s Counting Puppet task, which on the surface seems very similar. Le Corre's task discriminated subset-knowers from what we are calling high-numeral-knowers. Le Corre's participants knew that a character wanted, for example, six cookies. The puppet counted out 5 cookies and the child was asked "is that six?" Like our Last-Word task, the Counting Puppet task requires children to listen to an adult's standard (never tricky or unconventional) count, and to make a judgment about the result of that count. The main difference is that our Last-Word task uses the phrase "how many" in the test question, and this is enough to prompt many children to repeat the last counting word, whether or not they know that it represents the numerosity of the whole set. The Counting Puppet task, in contrast, can only be solved if the child realizes that the last word in a count (e.g., "five") represents that the set contains five individuals—i.e., if the child understands the cardinal principle.

Given that other studies have shown high within-child consistency across cardinality tasks including Give-N, the Counting Puppet task, What's-On-This-Card, etc. (Le Corre et al., 2006; Le Corre & Carey, in press; Wynn 1992), it seems fair to conclude that the Last-Word task *over*estimates cardinal-principle knowledge, rather than that Give-N and the other tasks all *under*estimate it. It would also seem prudent for future researchers wishing to assess children's cardinal-principle knowledge to avoid the phrase "how many."

Although the last-word rule is not the cardinal principle, it is worth studying in its own right. Three informative findings from the present study's Last-Word task were that (a) the great majority of children either got both trials correct or neither of two trials correct, indicating that they either knew the rule or didn't know it; (b) most children had learned the rule by the time they were "two"-knowers; (c) when children answered incorrectly, they either produced a different numeral or (less commonly) produced the numeral list itself (i.e., counted out loud).

These facts suggest that children's interpretation of the phrase "how many" changes as their understanding of counting grows. The earliest interpretation is probably that "how many" is a prompt to count (i.e., to produce the numeral list that they have memorized – see Fuson, 1988 for related findings). Our Last-Word results suggest that children soon figure out that "how many" is a question to be answered with a numeral, but many do not know how to decide which one. Most "two"-knowers and above even knew that "how many" should be answered with the last numeral used in counting—all this despite a preponderance of evidence showing that they do not understand how or why the last numeral used in counting denotes the numerosity of the set.

Finally, the present study explored children's knowledge of how counting implements the successor function, and related this understanding to other measures of mastery of the cardinal principle. The Direction and Unit tasks probed whether children knew that if you start with a set of "six" and add one, the resulting set has "seven," not "five" (the Direction task) and the resulting set has "seven," not "eight" (the Unit task). These results strongly support the hypothesis that high-numeral-knowers differ from subset-knowers in having worked out how counting represents natural numbers, as generated by the successor function. No-numeral-, “one”-, “two” and “three”-knowers utterly failed both the Direction and Unit tasks, and only high-numeral-knowers succeeded at both. “Four”-knowers succeeded at the Direction task but failed the Unit task.

The failure of the subset-knowers on both tasks is particularly telling. The Unit and Direction tasks were each arithmetic tasks—individual items were added or subtracted from sets and the child was given a forced-choice decision as to the cardinal value of the resulting set. Supporters of the principles-first position (Cordes & Gelman, 2005; Zur & Gelman, 2004) have argued that arithmetic tasks elicit the highest level of numerical understanding in preschool children, and that because Give-N is not an arithmetic task, it underestimates children's understanding of how counting represents number. However, the present study finds within-child consistency between Give-N and the Direction and Unit tasks, just as other studies have found within-child consistency between Give-N and other tests of cardinal-principle understanding.

The success of "four"-knowers on the Direction task might be thought of in two ways. First, some children classified as "four"-knowers on this task may well have had some fragile understanding of the cardinal principle. "Four"-knowers are relatively rare in the literature (e.g., of the 87 2- to 4-year-old children in LeCorre et al.'s (2006) studies, 8% were no-numeral-knowers, 15%, 18%, and 20% were "one"- "two"- and "three"-knowers respectively, and 32% were high-numeral/CP-knowers, but only 7% were "four"-knowers), suggesting that children are "four"-knowers for only a short time before completing their construction of the cardinal principle.

Second, working out how counting represents number is not accomplished in a single step. Children learn the exact cardinal meanings of "one," "two," "three" and "four" before they become high-numeral/CP-knowers, and the present studies show that they learn a rule that the last word of a count answers a "how many" question as well. Sarnecka and Gelman (2004) showed that subset-knowers understand that numerals depict *some* precise cardinal value, but that they don't know how to determine which value. It makes sense that children would work out the direction principle before the unit principle, for the latter presupposes the former. The present studies suggest that finally putting together the puzzle of how counting implements the successor function is indeed what turns a subset-knower into a high-numeral/CP-knower, but that “four”-knowers have almost all the pieces in place.

The present study has tried to characterize some of the partial knowledge children have as they figure out how counting implements the successor function, and more importantly to identify subcomponents of cardinal-principle knowledge itself. Characterizing children's knowledge at various points in this process will in turn constrain the theories we build, as we attempt to understand a remarkable intellectual achievement—children’s discovery of how the verbal numerals represent natural numbers.

References

Baroody, A. J., & Price, J. (1983). The development of the number-word sequence in the counting of three-year-olds. *Journal for Research in Mathematics Education, 14*, 361-368.

Briars, D., & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology, 20*, 607-618.

Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus,* 59-68.

Condry, K. F., & Spelke, E. S. (in press). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology – General.*

Cordes, S., & Gelman, R. (2005). The young numerical mind: When does it count? In J. I. D. Campbell (Ed.), *Handbook of Mathematical Cognition* (pp. 127-142). New York: Psychology Press.

Frye, D., Braisby, N., Lowe, J., Maroudas, C., & Nicholls, J. (1989). Young Children's Understanding of Counting and Cardinality. *Child Development, 60*(5), 1158-1171.

Fuson, K. C. (1988). *Children's counting and concepts of number*. New York: Springer-Verlag.

Fuson, K. C. (1992). Relationships between counting and cardinality from age 2 to age 8. In J. Bideaud, C. Meljac & J. Fischer (Eds.), *Pathways to number: Children's* *developing numerical abilities* (pp. 127-149). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Fuson, K. C., Richards, J., & Briars, D. J. (1982). The acquisition and elaboration of the number word sequence. In C. Brainerd (Ed.), *Progress in cognitive development research: Vol. 1. Children's logical and mathematical cognition* (pp. 33-92). New York: Springer-Verlag.

Fuson, K. C., Secada, W. G., & Hall, J. W. (1983). Matching, counting, and conservation of numerical equivalence. *Child Development, 54*(1), 91-97.

Gelman, R. (1993). A rational-constructivist account of early learning about numbers and objects. In D. L. Medin (Ed.), *The psychology of learning and motivation. Advances in research theory* (pp. 61-96). San Diego: Academic Press.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Greeno, J. G., Riley, M. S., & Gelman, R. (1984). Conceptual competence and children's counting. *Cognitive Psychology, 16*, 94-143.

Hurford, J. (1987). *Language and number : The emergence of a cognitive system*. New York: B. Blackwell.

Klahr, D. (1984). Transition Processes in Quantitative Development. In R. Sternberg (Ed.), *Mechanisms of Cognitive Development* (pp. 101-139). San Francisco: W. H. Freeman.

Klahr, D., & Wallace, J. G. (1976). *Cognitive development: An information-processing view*. Hillsdale, NJ: Erlbaum.

Le Corre, M., & Carey, S. (in press). One, two, three, four, nothing more: How numerals are mapped onto core knowledge of number in the construction of the counting principles. *Cognition.*

Le Corre, M., Li, P., & Jia, G. (2003). On the role of singular/plural in number word learning. *Society for Research in Child Development,* Tampa, FL.

Le Corre, M., Van de Walle, G., Brannon, E. M., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology, 52*(2), 130-169.

Li, P., Le Corre, M., Shui, R., Jia, G., & Carey, S. (2003). Effects of plural syntax on number word learning: A cross-linguistic study. *28th Boston University Conference on Language Development,* Boston, MA.

Markman, E. M. (1979). Classes and collections: Conceptual organizations and numerical abilities. *Cognitive Psychology, 11*, 395-411.

Miller, K. F., Smith, C. M., Zhu, J. J., & Zhang, H. C. (1995). Preschool origins of cross-national differences in mathematical competence -- The role of number-naming systems. *Psychological Science, 6*(1), 56-60.

Miller, K. F., & Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Cognitive Development, 2*, 279-305.

Rittle-Johnson, B., & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review of the literature. In C. Donlan (Ed.), *The development of mathematical skills* (pp. 75-110). Hove, England: Psychology Press.

Sarnecka, B. W., & Gelman, S. A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition, 92*, 329-352.

Sarnecka, B. W., Kamenskaya, V. G., Yamana, Y., Ogura, T., & Yudovina, J. B. (in press). From grammatical number to exact numbers: Early meanings of “one," “two,” and “three” in English, Russian, and Japanese. *Cognitive Psychology.*

Schaeffer, B., Eggleston, V. H., & Scott, J. L. (1974). Number development in young children. *Cognitive Psychology, 6*, 357-379.

Wagner, S. H., & Walters, J. (1982). A longitudinal analysis of early number concepts. In G. Foreman (Ed.), *Action and thought: From sensorimotor schemes to symbolic operations* (pp. 137-161). New York: Academic.

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*, 155-193.

Wynn, K. (1992). Children's acquisition of number words and the counting system. *Cognitive Psychology, 24*, 220-251.

Zur, O., & Gelman, R. (2004). Young children can add and subtract by predicting and checking. *Early Childhood Research Quarterly, 19*, 121-137.