Developmental Psychology
2018, Vol. 54, No. 5, 875–889


0012-1649/18/$12.00 http://dx.doi.org/10.1037/dev0000469

# How High Can You Count? Probing the Limits of Children's Counting

Tasha Posid
The Ohio State University

Sara Cordes
Boston College

While much research has focused on understanding the process by which young children learn to count, little work has explored the effects of direct instruction on this process. In the current study, we explored the impacts of training children in an explicit counting procedure on two distinct cardinality tasks. Two- to 5-year-old children first participated in a Give-N task in which counting proficiency was assessed, and then participated in a short instruction session where explicit counting was modeled and encouraged. Following training, children were significantly better at identifying which of two cards contained a set size outside of their range of counting mastery (Huang, Spelke, & Snedeker, 2010) and were more likely to improve on a secondary numerical production task (Give-N; Wynn, 1990, 1992) compared with children in the control group. Not surprisingly, a greater proportion of children in the count training condition overtly counted during the cardinality task, a strategy that was found to be the strongest indicator of performance. Together, results reveal that even 5 min of counting instruction greatly increases the likelihood that a child will engage in counting behavior and results in improvements in cardinality judgments in two distinct numerical tasks.

*Keywords:* counting, number concepts, cardinality, numerical cognition



Just as young children can recite the alphabet before they can read or write, young children can count aloud before they understand what those number words really mean (e.g., Condry & Spelke, 2008; Gallistel & Gelman, 1992; Wynn, 1990). In time, however, children come to understand the meaning behind those number words that they can recite out loud and more importantly, they learn that reciting those words in sequence (i.e., counting) constitutes a reliable strategy for determining the cardinality of a set ("the Cardinal Principle;" Gelman & Gallistel, 1978). Evidence suggests that counting fluency is critical for later achievement in the classroom, with young children's mastery of the counting principles predictive of their foundational symbolic number understanding (Miller & Gelman, 1983), as well as later arithmetic abilities and achievement scores (Duncan et al., 2007; Geary, 2011; Stock, Desoete, & Roeyers, 2009). Thus, understanding these early precursors such as the count acquisition process are important to understanding foundational mathematical abilities that impact numerical understanding and proficiency even late into development. While much research has focused on understanding the process by which young children learn to count (e.g., Gelman & Gallistel, 1978; Le Corre & Carey, 2007; Miller & Stigler, 1987), little work has explored the effects of direct instruction on this process (except see Mix, Sandhofer, Moore, & Russell, 2012; Petersen et al., 2014). The current study aims to explore whether brief training on explicit counting procedures can have an immediate impact on children's cardinality judgments.

## Background

The count acquisition process has been extensively studied. Many researchers have characterized the acquisition of counting in the preschool years as a slow, stage-like process in which children first learn the meanings of the first three- to four-number words exactly and sequentially (termed the "Discontinuity Hypothesis;" e.g., Carey, 2001, 2009; Le Corre & Carey, 2007; Le Corre, Van de Walle, Brannon, & Carey, 2006; Sarnecka & Lee, 2009). Evidence for this comes from the Give-N task, in which children are asked to give an experimenter a certain number of toys from a pile of toys (Wynn, 1990, 1992). Children who can produce one item, but no more, are called "one-knowers," children who can produce one and two items, but no more, are called "two-knowers," and so on. Thus, children progress through "knower-levels," beginning as subset-knowers—where they understand "one," then "two," then "three," and so forth. Then, it is only after acquiring an understanding of the first few numbers (up to three or four) that children come to understand that the purpose of counting is to determine the cardinality of a set (acquiring the Cardinal Principle and becoming "CP-knowers"), which allows them to then produce sets as large as they can count (Le Corre & Carey, 2007; Le Corre et al., 2006; Lee & Sarnecka, 2010; Sarnecka & Lee, 2009; Slusser & Sarnecka, 2011). It is only at this point

This article was published Online First March 8, 2018.

Tasha Posid, Department of Psychology, The Ohio State University; Sara Cordes, Department of Psychology, Boston College.

We thank the members of the Infant and Child Cognition Lab who assisted in data collection and coding. We would also like to thank the Boston Children's Museum and *Living Laboratory* at the Museum of Science, Boston. Finally, we thank Ellen Winner and Elida Laski for their thoughtful feedback on a previous version of this paper. Funding was provided by a NSF CAREER award 1056726 to S.C.

Correspondence concerning this article should be addressed to Tasha Posid, Department of Psychology, The Ohio State University, 268 Psychology Building, 1835 Neil Avenue, Columbus, OH 43210. E-mail: posid.1@osu.edu

that children demonstrate a flexibility with their understanding of the number words, applying them to novel sets and novel situations (e.g., Bar-David et al., 2009; Slusser & Sarnecka, 2011). Thus, according to the Discontinuity Hypothesis, early counters do not have the requisite knowledge available to succeed on cardinality tasks, even when counting is supported.

Alternatively, other researchers argue that children (including those classified as "subset knowers" on the Give-N task) have an inherent understanding of the cardinality principle and the count procedure (termed the "Continuity Hypothesis;" e.g., Cordes & Gelman, 2005; Gelman & Gallistel, 1978). Early poor performance on cardinality tasks is thus not due to a failure to understand how counting works, but instead due to difficulties in carrying out a long and complicated procedure which taxes working memory (i.e., a procedure which requires remembering *to* count, remembering a list of words [the count list, remembering which items were already counted and which are yet to be counted, etc.). Thus, according to the Continuity Hypothesis, it is not that young children do not understand the counting principles, but instead are hindered by performance issues, which limit how well they can engage in cardinality tasks. As such, those classified as Subset-Knowers on the Give-N task could demonstrate an understanding of cardinal number long before successfully producing a set of six items.

In support of the Continuity Hypothesis, evidence suggests that young children may have a precocious understanding of how number words work, understanding some of the principles behind cardinality labeling long before being able to correctly apply labels to these large sets. For example, Lipton and Spelke (2006) examined whether 5-year-old children understood the meaning of words for large numbers outside of their verbal count list (e.g., if a child could only count to 20, they might be asked about 40). They found that young children with limited counting abilities understood the logic of how number words are applied to sets such that for example, when a set with "forty" items has an item removed, the set no longer has "forty" items. Their results demonstrate that children understand symbolic number concepts even before they master the verbal count routine (Lipton & Spelke, 2006). Similarly, Condry and Spelke (2008) examined subset-knowers' errors when labeling the cardinality of sets outside of their range of mastery and found that although one- and two-knowers were unable to appropriately label the cardinality of a large set (e.g., "eight"), they already understood that a specific numerosity is singularly labeled (e.g., a set containing "eight" cannot also contain "four").

Other data, however, suggest that this precocious understanding may instead be indicative of an inappropriately overextended understanding of the concepts to which number words (outside their range of mastery) refer. That is, evidence suggests that less-than-proficient counters expect number words to not only refer to numerical attributes, but also perceptual attributes of a set. This overextended understanding results in a hesitancy to use the same number word to refer to multiple sets, despite identical cardinalities. For example, Slusser and Sarnecka (2011) labeled the cardinality of a set (e.g., "This set has *eight* turtles") and then asked children to pick another set (of two options) containing the same specific numerosity. Surprisingly, it was not until children reached the level of CP-knower that they readily (and correctly) applied the same number word to two different sets. Despite having these number words within their count list, less-skilled counters did not appear to understand that these number words (e.g., "eight") referred to set cardinality alone. In an even more striking demonstration, Huang, Spelke, and Snedeker (2010) used a labeling paradigm to teach two- and three-knowers the next number in their verbal count list that they did not yet conceptually understand. For example, two-knowers were shown pairs of cards containing sets of animals and the cards were labeled as containing "three dogs" and "*not* three dogs" (similarly, three-knowers were shown "four dogs" and "*not* four dogs"). Results demonstrated that two-knowers were very limited in their ability to generalize this new number knowledge, instead relying upon non-numerical features of the training material (e.g., the specific animal on the card). That is, two-knowers were able to correctly identify the card containing three items when the category of items on the training cards matched that used in test (e.g., they were trained to identify sets of three dogs, and then tested with different sets of dogs), but they performed at chance levels when basic level category information varied across trials (e.g., they saw dogs in training and elephants in test, Experiment 3). In contrast, three-knowers performed well-above chance for all comparisons, suggesting that these children had a much richer understanding of number. Based on their findings, the authors concluded that early counters fail to map newly learned number words of their verbal count list to full-fledged, abstract numerical concepts.

Evidence suggests that these paradigms in which sets were labeled as containing "three" and "*not* three" items (or "four" and "*not* four;" Huang et al., 2010), may not have been sufficient for promoting conceptual understanding in young children. Early language learners have been shown to typically construe labels as referring to categories (i.e., "dog"), not set characteristics (Waxman & Booth, 2001); thus, it is not surprising that children did not expect novel number words to refer to purely numerical concepts. Instead, evidence from training studies highlight the importance of both labeling *and* counting sets for young children to promote performance on cardinality tasks. For example, Petersen et al. (2014) found that children who participated in a 5-week picture training paradigm, in which sets were counted and labeled for them, demonstrated improved numerical understanding. Relatedly, Mix, Sandhofer, Moore, and Russell (2012) report that 3.5-year-old children who participated in 6 weeks of training—in which experimenters counted *and* labeled the numerosity of sets depicted in picture books—showed marked improvement on the Give-N task, with improvements apparent after only 3 weeks. In contrast, children who participated in training which exclusively involved counting (but no labeling) or labeling (but no counting; similar to Huang et al., 2010), showed no improvements, even after the full 6 weeks of training. Critically, these findings demonstrate the importance of both counting sets *and* labeling cardinal values for children's developing sense of cardinality. However, much more work is needed to explore the mechanism responsible for these changes in young children's performance on cardinality tasks following extended training. Was it simply the adult modeling of counting and labeling that promoted children's numerical performance? That is, is it sufficient for a child to observe counting and labeling behavior, or must the child actually engage in counting themselves in order to promote success on numerical tasks? These questions have important theoretical and practical implications.

Unlike children's numerical knowledge, much less is known about *when* children choose to engage in counting behaviors. Evidence suggests that some children spontaneously count when engaging in number tasks (Gelman & Tucker, 1975; Zur & Gelman, 2004), and that this tendency increases with the child's knower-level and with the set sizes presented (Le Corre & Carey, 2007; Le Corre, Van de Walle, Brannon, & Carey, 2006; Sarnecka & Carey, 2008). For example, Sarnecka and Carey (2008) reported that only 16% of subset-knowers spontaneously counted in their Give-N task, compared with 67% of CP-knowers. However, whether the child's counting behavior predicted performance on their numerical tasks was not addressed. Evidence does suggest that children who engage in counting out-perform those who do not on a variety of tasks. Fuson, Secada, and Hall, (1983) found that 5-year-old children who were told to use counting or matching to solve a Piagetian conservation-of-number task out-performed children who received no instruction. Bar-David et al. (2009) asked children to retrieve "just enough" socks for caterpillars with a varying number of feet. They found that for subset-knowers, the greatest predictor of success in their task was whether the child chose to engage in verbal counting. Similarly, Lipton and Spelke (2005) and Wylie, Jordan, and Mulhern (2012) found that older children who used explicit counting were more likely to succeed on numerical calculation tasks and demonstrated a richer, more abstract understanding of the relationship between number words and their associated numerosity. Thus, it appears that a child's engagement in explicit counting behaviors is beneficial to numerical performance (see Bar-David et al., 2009; Fuson et al., 1983; Gelman & Tucker, 1975; Le Corre & Carey, 2007; Lipton & Spelke, 2005; Sarnecka & Carey, 2008; Wylie et al., 2012; Zur & Gelman, 2004). If this is the case, then is it possible to encourage counting behavior in children as a means of promoting success on cardinality tasks that would otherwise prove too challenging?

The answer to this question has serious theoretical implications for prominent accounts of the acquisition of early counting. According to the Continuity Hypothesis (e.g., Cordes & Gelman, 2005; Gelman & Gallistel, 1978), young children have an inherent understanding of the purpose of counting, yet are simply limited in their ability to carry out the complicated procedure. Thus, if proper counting is modeled and encouraged, then children should already have the prerequisite conceptual understanding to benefit from counting. As such, promoting appropriate counting behavior in young children should lie within their Zone of Proximal Development (ZPD; Vygotsky, 1978), resulting in performance improvements on cardinality tasks. On the other hand, the Discontinuity Hypothesis (Carey, 2001, 2009; Le Corre & Carey, 2007; Le Corre et al., 2006; Sarnecka & Lee, 2009) posits that children's performance on cardinality tasks is limited by a lack of conceptual understanding of how counting works; therefore, training subset-knowers to count before they have acquired requisite cardinal knowledge should have little-to-no impact on their performance on cardinality tasks. That is, subset-knowers do not understand how counting works and so counting is a meaningless procedure unlikely to aid their performance in cardinality tasks.

In the present study, we examined how readily subset-knowers can be encouraged to use counting to identify novel set sizes and whether their use of counting leads to greater success on two different number tasks, compared to nontrained controls. In line with Mix et al. (2012), we explicitly trained young children with both counting and labeling of sets. However, in contrast to studies employing drawn-out training periods (i.e., 6 weeks in Mix et al., 2012; 5 weeks in Petersen et al., 2014), our training consisted of a single, 5-min training session. This novel, single-session paradigm stands in contrast to previous studies examining the effects of direct instruction by testing the *immediate* impacts of a brief instruction. Specifically, this paradigm mirrors the types of activities that could be used in the real-world, from in-class instruction (by teachers or peers), to at-home parent–child play activities or even electronic applications. We used a modified training paradigm (adopted from Huang et al., 2010), in which we modeled both counting and labeling the cardinality of sets in order to identify a target number of items in a card choice task. In Experiment 1, subset-knowers (two-, three-, four-, and five-knowers) were trained to identify sets containing six items—a number outside of their current range of mastery—to explore how readily children acquire novel number concepts. We then examined if children spontaneously generalize the benefits of this counting procedure to a different numerical task examining cardinal understanding (a second Give-N task). In Experiment 2, we continued this line of investigation, looking exclusively at two-knowers identifying sets of three items (as in Huang et al., 2010). Importantly, in secondary analyses in both experiments, we observed children's counting behavior in order to determine if proper use of the counting procedures fostered greater success in these tasks. We asked: (a) If encouraged to count, can children learn to identify a set size outside of their range of mastery? (b) Do the benefits of count training (brief instruction) generalize to other cardinality tasks (that differ substantially in nature and task demands)? (c) Does this encouragement impact their propensity to count? and (d) Does a child's counting behavior mediate any observed improvements on these numerical tasks?

## Experiment 1

### Method

**Participants.** Two- to 5-year old children were initially prescreened using the Give-N task (see below) to determine their knower-level. Because knower-level could not be randomly assigned, inclusion in the study was determined by knower-level, such that once a full sample of children at a particular knower-level was obtained, additional children who were screened and classified as falling within that knower-level were not run in the full procedure. For example, once a full sample of two-knowers had participated in the study, additional children who were classified as two-knowers were not included and instead participated in other ongoing studies in the lab. Therefore, a significantly greater number of children were screened in the Give-N task than were included in this study.

In the process of screening children, a small (yet notable) number of children were classified as five-knowers,[1] as has been found by many other labs (Gunderson, Spaepen, & Levine, 2015; Mussolin, Nys, Content, & Leybaert, 2014; Nikoloska, 2009; Pinhas, Donohue, Woldorff, & Brannon, 2014; Wagner & Johnson, 2011). Although the proportion of five-knowers was significantly less than that of two-, three-, or four-knowers, we continued to screen children until we obtained a large enough sample of five-knowers to include in the study because we thought it to be both theoretically and practically important to characterize this group of children.

In total, 186 2- to 5-year-old children were included in the final sample of participants who were run through the entire procedure. Based on their knower-level, children were randomly assigned to the Counting ($N = 93$) or to the Control ($N = 93$) conditions within each knower-level (see Table 1). Children classified as one-knowers were excluded from this study because pilot testing indicated that one-knowers possessed such little number knowledge as to not understand the rules of the game.[2] Also, children classified as CP-knowers (who could produce six items in the Give-N task) were also excluded, as the goal was to explore the impacts of counting training in children who presumably did not already have an understanding of the Cardinal Principle (i.e., subset-knowers). An additional 13 children were excluded due to their failure to complete all 14 test trials (eight), for parental interference (one), or for scoring $<$50% accuracy on known-number trials (four). Children took part in the study at one of two local museums (the Boston Children's Museum or the Museum of Science, Boston) or during one visit to our laboratory. For those who visited the lab, children's names were obtained via local birth records. There were no general differences in demographics or task performance across children run at the three research sites.

**Materials.**

***Give-N task.*** Fifteen 2-in. yellow rubber ducks were used in the first Give-N task and 15 1.5-in. vinyl green frogs were used in the second Give-N task. A 5.25-in. × 2.25-in. circular blue plastic basket was used in both tasks as the "pond" in which children were asked to place the toy animals.

***Card task.*** Stimuli consisted of laminated cards, each 8.5 × 5.5 in. On each trial, children were presented with a pair of cards, each depicting an array of animals. For both training and test, within each card pair, the same animal was depicted (e.g., chickens vs. chickens), but across card pairs, different animals were used such that no animal was repeated for more than one pair of cards during training or test (e.g., dogs, pigs, horses, etc.). The arrays of animals pictured on the cards varied in their spatial arrangement (e.g., vertical rows, horizontal rows, triangles), and the side of the target card and the side of the larger quantity was varied across trials.

In training, each card pair contained one card with the target number of animals (six), and the other card depicted one, two, four, five, eight, or 12 animals, for a total of six training trials. In training, the animals on each pair of cards were of equal size, thus continuous variables such as cumulative area and contour were correlated with number.

In test, children saw cards depicting six animals paired with cards depicting one, two, three, four, nine, 12, and 18 animals. Each numerical quantity appeared twice in test, for a total of 14 trials. While the size of the animals pictured on each pair of cards varied from pair to pair, items within each pair were approximately matched for cumulative area in test so as to discourage the use of non-numerical cues for responding.

**Procedure.** All children participated in three tasks in the following sequence: (a) Give-N task; (b) Card task; (c) Give-N task again.

***Give-N.*** [3]Children completed the Give-N task (modeled after Wynn, 1990, 1992) in order to determine their knower-level (Le Corre & Carey, 2007). Fifteen small yellow ducks were placed in a pile in front of the child next to a blue basket. Children were

[1] According to the discontinuity account (e.g., Carey, 2009), subset knowers' understandings of the first few count words are derived from a mapping formed between those words and an underlying parallel individuation system (aka the object-file system) used for tracking objects in the world, a system which is only capable of tracking as many as four objects at a time. Once children acquire an understanding of the first few count words (up to "three" or "four") and subsequently acquire the Cardinal Principle, these words are then spontaneously mapped to a distinct system of number representation, the Approximate Number System (ANS). Thus, children progress from being three- or four-knowers, reliant upon the parallel individuation system, to Cardinal Principle-knowers, capable of representing all real numbers (Carey, 2009; Le Corre & Carey, 2007; Le Corre et al., 2006; Sarnecka & Gelman, 2004; Wynn, 1990, 1992). Importantly, it should be impossible to identify children as five-knowers, because the parallel individuation system is thought to be incapable of tracking as many as five items simultaneously (Feigenson, Dehaene, & Spelke, 2004). On the other hand, continuity account theorists have posited that children's number word knowledge is not initially mapped to the parallel individuation system, such that early number knowledge is not constrained by the three to four item set size limit (Cordes & Gelman, 2005; Gelman & Gallistel, 1978; Wagner & Johnson, 2011). In line with this account, a number of studies have identified a subset of children who are best classified as five-knowers using the Give-N task and similar criteria to Le Corre and Carey (2007; see Gunderson, Spaepen, Gibson, Goldin-Meadow, & Levine, 2015; Mussolin et al., 2014; Nikoloska, 2009; Piantadosi, Jara-Ettinger, & Gibson, 2014; Pinhas et al., 2014; Wagner & Johnson, 2011). Similarly, we also identified a subset of children as five-knowers in our task who successfully and reliably produced sets of five (but not six) items when requested.

[2] Although it may have been theoretically interesting to include one-knowers in the present study, pilot testing revealed that these children did not understand task demands and typically became inattentive during our card task and were unable to complete the tasks in the study.

[3] Although our Give-N task was modeled after that of Le Corre and Carey (2007), some differences in how the task was administered may have impacted classification of knower-levels. For example, after producing a set of N, children in our task were not asked "Is that N?" and allowed to fix their responses. Because children were not given the opportunity to change their answers, it is possible that the knower-level classifications of a few children may have been underestimated. Importantly, numerous other studies have also excluded the "Is that N?" question following the child's numerical production, with comparable findings to those studies including the question (e.g., Fluck & Henderson, 1996; Gunderson, Spaepen, Gibson, Goldin-Meadow, & Levine, 2015; Huang et al., 2010; Kaminski, 2015; Mix et al., 2012; Munn & Stephen, 1993; Opfer, Thompson, & Furlong, 2010 (Exp. 3); Negen & Sarnecka, 2012; Posid & Cordes, 2014; Schaeffer, Eggleston, & Scott, 1974; Wagner & Johnson, 2011). Moreover, we found the rate of spontaneous counting during our initial Give-N task (11.5% of children in a random sample of 52 videos coded post hoc) to not differ significantly from that reported by Sarnecka and Carey (2008) who included the "Is that N?" question in their procedure (16%); $\chi^2(1, N = 101) = 0.31$, $p > .5$. Most importantly, however, the Give-N tasks administered in our study (both at the beginning of the session and at the end of the session, and across the two different conditions) were all administered in exactly the same manner, ensuring that our pattern of results across the two conditions could not be attributed to slight procedural differences in the administration of the Give-N task.

Table 1
*Distribution of Participants Across Knower-Levels and Conditions in Experiment 1*

| | Control Condition | Counting Condition |
|---|---|---|
| 2-knowers: | $N$ = 25 ($M$ = 34.8 mos, $SD$ = 6.57 mos) | $N$ = 26 ($M$ = 35.8 mos, $SD$ = 9.15 mos) |
| 3-knowers: | $N$ = 25 ($M$ = 36.4 mos, $SD$ = 5.37 mos) | $N$ = 22 ($M$ = 38.0 mos, $SD$ = 6.47 mos) |
| 4-knowers: | $N$ = 22 ($M$ = 43.4 mos, $SD$ = 7.38 mos) | $N$ = 24 ($M$ = 44.3 mos, $SD$ = 8.08 mos) |
| 5-knowers: | $N$ = 22 ($M$ = 43.4 mos, $SD$ = 7.38 mos) | $N$ = 21 ($M$ = 44.3 mos, $SD$ = 7.82 mos) |

asked to put varying numbers of ducks, from one to six, into a blue basket ("Can you make *N* ducks jump into the pond?"). The experimenter first asked the child to place "one duck" into the pond. If the child successfully did so, then the experimenter would continue and ask for "two ducks" and so forth. If the child failed to give the correct quantity (*N*), the experimenter then asked for *N*-1 ducks again. If the child successfully gave *N*-1 ducks, the experimenter then asked for *N* ducks again. If the child successfully gave *N* ducks, the experimenter continued to ask for higher quantities. If the child failed to give *N* ducks for the second time, the experimenter ended the game. Thus, a child had to fail to produce a set of *N* items twice, *or* to produce a set of six items twice correctly, in order for the game to end. Children were assigned knower-levels based on the highest number of items they produced correctly at least 67% of the time (e.g., two-knowers could produce two ducks but not three; three-knowers could produce three ducks but not four; and so on). If the child correctly gave every quantity through six, the child was classified as a CP-knower.

***Card task: Familiarization for counting condition.*** [4]For children randomly assigned to the Counting condition, participants were presented with a single card depicting the target number of animals (six). The experimenter labeled the number of animals on the card and demonstrated how to count the items on the card by pointing to each individual item while tagging it with the corresponding count word (e.g., "This card has *six* animals on it! See, one, two, three, four, five, six!"). This was repeated for a different card. After two single presentations, children were then shown two cards at a time, with one card containing the target number of animals and the other card containing a different number. Again, the experimenter demonstrated how to count each item on the card to find the target number of animals (e.g., "This card has *six* animals on it! See, one, two, three, four, five, six! But this [other] card does not have six animals on it. See, one, two, three, four . . ."). The experimenter always counted in a consistent manner (e.g., the experimenter began counting the top row of items, counting from left to right, then moved to the second row and counted from left to right, etc.). Unlike Huang et al. (2010), the experimenter did not refer to the types of animals when counting arrays of card pairs and instead always used the term "animals" when referring to the number of items on each card. From set-to-set, a variety of animal species were presented across training and test and children did not see the same animal more than once. Participants did not move to the test phase of the Card task until they had participated actively on every familiarization trial, either by correctly counting the items on each card or by correctly selecting the target card from the pair presented. If a child did not use correct, one-to-one counting during this training phase or if a child selected the distractor card when asked to select the target value, they were given corrective feedback by the experimenter.

***Familiarization for control condition.*** For children randomly assigned to the Control condition, familiarization trials were identical to the Counting condition, except that the experimenter never counted the items on the card, and instead said "This card has *six* animals on it! But this [other] card does not have six animals on it!" as they simultaneously pointed to the card, then circled the entire set in a single motion. Again, the experimenter did not refer to the types of animals when counting arrays of card pairs; however, a variety of animal species were presented across training and test and children did not see the same animal more than once. To equate the longer length of time children were exposed to each card in the Counting condition, children were shown the same training card pairs a second time in the Control condition and were asked to select which of the two cards contained the target number. Participants did not move to the test phase of the Card task until they had participated actively in training by correctly selecting the target card from each pair. Again, if a child selected the distractor card when asked to select the target value during familiarization, corrective feedback was given.

***Card task: Test.*** During test, all children saw the same 14 novel card pairs and were asked to select the card containing the target number of items (e.g., "Which card has *six* animals on it?"). Children were always asked to select the target number (six). The experimenter provided positive, neutral feedback regardless of the child's choice ("Great job!" or "Thank you!") and recorded the child's responses. As in Huang et al. (2010), children were presented with comparisons of both smaller, known-numbers and larger, unknown-numbers.

***Second Give-N.*** After children completed the Card task, they participated in a second Give-N task, identical in procedure to the first Give-N task, this time involving toy frogs jumping into a pond (blue basket).

**Data coding and analyses.** During test, the experimenter recorded the child's response and also noted whether or not the child consistently attempted to use explicit, verbal counting to find the target card (as per Bar-David et al., 2009). Children who spontaneously attempted to engage in explicit, verbal counting for the majority of test trials (regardless of whether the items were correctly counted or not) were coded as "Counters" and children who did not engage in explicit verbal counting were coded as "Non-Counters." Two independent observers coded children's counting strategies for a subset of the participants (35%) and intercoder

[4] It should be noted that while this condition is called the counting condition, the cardinality of the set was both labeled *and* counted for the children during training. In contrast, in the control condition, counting was not modeled, but the cardinality of the set was still labeled for the children. Thus, the "counting" and "control" conditions would be more accurately referred to as the "label and count" and "label only" conditions, but in the interest of space and clarity, we have used "counting" and "control."

reliability was 93.0%. Any disagreements were settled by a third-party rater, and children received a designation of "Counter" or "Non-Counter" based on the agreement of two of the three raters.

For purposes of data analyses, accuracy was computed separately for "known-number trials" versus "unknown-number trials" for each knower-level (as per Huang et al., 2010). Because number knowledge necessarily varies as a function of knower-level, different numbers of trials contributed to the computation of known and unknown accuracy scores for each knower-level. For example, whereas known-number trials for two-knowers included those trials in which a card with 6 items was pitted against a card with one or two animals (and unknown-number trials included six vs. three, four, nine, 12, 18), known-number trials for three-knowers included trials pitting six against one, two, and three items (with unknown-number trials including six vs. four, nine, 12, and 18), and so forth.

Preliminary analyses found no main effects or interactions with gender ($p > .05$), and thus this variable was not included in further analyses.

## Results

### Performance on the Card Task

**Known-number trials.** A Condition (Counting vs. Control) × Knower-Level (two-, three-, four-, and five-knowers) univariate ANOVA on percent correct for known-number trials revealed a near-significant main effect of knower-level, $F(3, 178) = 2.22$, $p = .087$, $\eta_p^2 = .036$, such that children's accuracy improved as knower-level increased. Children of all knower-levels were proficient at these easier known-number trials, with even the least proficient knower-levels performing well above chance-level (chance = 50%; all $p$'s < .001, all *Cohen's d′*s > 5.05; see Table 2). Although children in both the Counting and Control conditions performed well above chance level (Counting condition: $M = 93.8\%$, $t(92) = 27.5$, $p < .001$, *Cohen's d* = 5.73; Control condition: $M = 88.2\%$, $t(92) = 35.6$, $p < .001$, *Cohen's d* = 7.42), a main effect of condition revealed that children in the counting condition performed significantly more accurately than their peers in the control condition, $F(1, 178) = 9.28$, $p = .003$, $\eta_p^2 = .05$. Moreover, the Condition × Knower-Level interaction was not significant, $F(3, 178) = 0.346$, $p > .7$, $\eta_p^2 = .006$, indicating that this pattern held across all knower-levels (see Figure 1).[5]

**Unknown-number trials.** A Condition (two) × Knower-Level (four) univariate ANOVA was conducted on accuracy on unknown-number trials. Results revealed a main effect of knower-level ($F(3, 178) = 11.7$, $p < .001$, $\eta_p^2 = .164$), such that 2- and 3-knowers performed comparably ($p > .9$, *Cohen's d* = .012), but both groups performed less accurately than 4-knowers ($p = .007$, *Cohen's d* = .54), who, in turn, were less accurate than five-knowers ($p = .022$, Cohen's $d = .39$). Importantly, there was also a main effect of condition, $F(1, 178) = 7.02$, $p = .009$, $\eta_p^2 = .038$, such that although children in both conditions performed above chance level (Control condition: $t(92) = 2.98$, $p = .004$, Cohen's $d = .62$; Counting condition: $t(92) = 6.13$, $p < .001$, Cohen's $d = 1.28$), children in the Counting condition ($M = 68.7\%$) were significantly more accurate on unknown-number trials than children in the Control condition ($M = 58.9\%$). However, this main effect was qualified by a significant interaction between knower-level and condition, $F(3, 178) = 5.61$, $p = .001$, $\eta_p^2 = .086$ (see

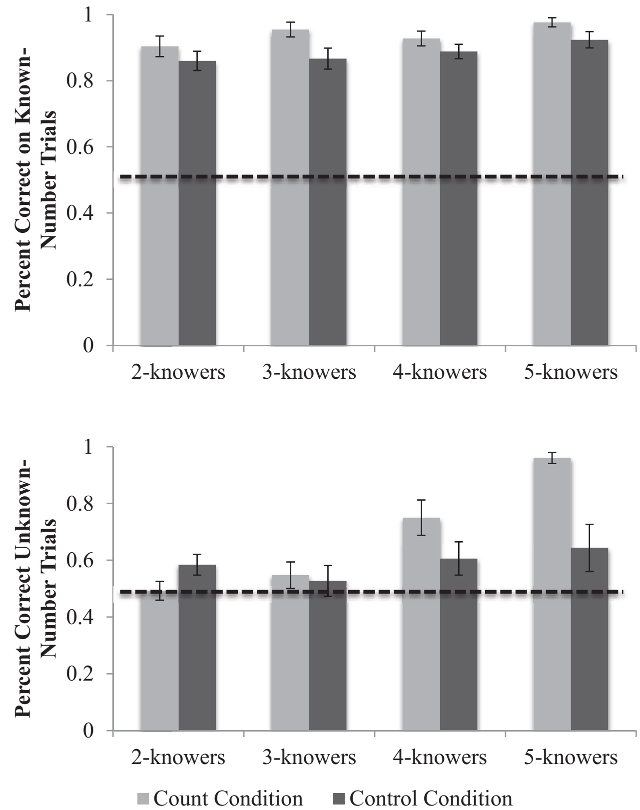


*Figure 1.* Children's performance on known-number trials (top) and unknown-number trials (bottom) of the Card task in Experiment 1 as a function of knower-level and condition. The black dashed line indicates chance-level (50%) performance. Error bars reflect standard error of the mean.

Table 2
*Mean Accuracy on Known- and Unknown-Number Trials in the Card Task in Experiment 1*

| | Known-number trials: | | Unknown-number trials: | |
|---|---|---|---|---|
| | Control | Counting | Control | Counting |
| 2-knowers: | 86.0% | 90.4%*** | 58.4% | 49.2% |
| 3-knowers: | 86.7% | 95.5%*** | 52.5% | 54.5% |
| 4-knowers: | 88.7% | 92.7%*** | 60.6% | 75.0% |
| 5-knowers: | 92.3% | 97.6%*** | 64.3% | 96.0%*** |

*** $p < .001$.

[5] Although five-knowers appeared to show the greatest improvement in the counting condition, this main effect of condition held even when five-knowers were excluded from analyses. For known-number trials, a Condition (two: Counting vs. Control) × Knower-Level three: two-, three-, and four-knowers) was rerun on percent correct. Results indicate no main effect knower-level, $F(2, 138) = .664$, $p > .5$ and a main effect of condition, $F(1, 138) = 6.6$, $p = .011$, $\eta_p^2 = 046$; no interaction, $F(2, 138) = .456$, $p > .6$.

Figure 1). Follow-up analyses revealed that the impact of encouraging children to count varied as a function of knower-level. Surprisingly, two-knowers performed slightly better in the control condition than the counting condition, $t(2) = 1.85$, $p = .07$, Cohen's $d = .52$. Three-knowers demonstrated comparable performance on unknown-number trials across the two conditions, $t(45) = 0.281$, $p > .7$, Cohen's $d = .08$. In contrast, both four-knowers (marginally) and five-knowers performed more accurately in the counting condition compared with the control condition (four-knowers: $t(44) = 1.67$, $p = .102$, Cohen's $d = .49$; five-knowers: $t(40) = 3.72$, $p = .001$, Cohen's $d = 1.15$; see Table 2).[6]

Additionally, a follow-up one-way ANOVA examining the effects of knower level on percent correct on unknown-number trials in exclusively the control condition indicated no main effect of knower level, $F(1, 89) = .711$, $p > .5$, $\eta_p^2 = .023$, confirming that there were no performance differences across knower levels, with only two-knowers (inexplicably) performing significantly above chance levels (two-knowers: $t(24) = -2.28$, $p = .032$, Cohen's $d = .93$; all other groups: $p$'s $> .05$). In contrast, a one-way ANOVA on unknown trials in the counting condition revealed that when children were encouraged to count, accuracy on unknown-number trials increased as a function of knower-level, $F(3, 89) = 22.5$, $p < .001$, $\eta_p^2 = .431$, with performance exceeding chance levels for children classified in the highest knower-levels (two- and three-knowers: $p$'s $> .6$; four- and five-knowers: $p$'s $< .001$).

**Second Give-N performance.** We also assessed the impact of our counting manipulation on children's performance on our follow-up cardinality task: the second Give-N task (conducted *after* the Card task). Performance on this second Give-N task was compared with performance on the initial Give-N task (used to assess knower-level). Overall, 32 of 186 (17.2%) children improved on the second Give-N task by correctly producing a larger set size than in the first Give-N task. Importantly, the vast majority of improvements were observed in children who participated in the counting condition, with 27 of the 32 children having participated in the counting condition, $\chi^2(1, N = 186) = 18.3$, $p < .001$, *Cramer's V* $= .313$ (see Figure 2).[7] This difference in overall improvements held across knower-levels, such that a greater proportion of children improved on the second Give-N in the counting condition than in the control condition for all four knower-levels (see Table 3).

Moreover, although most improvements involved children producing a set that was one item larger than in the first Give-N task (75%), a non-negligible proportion of children improved by two or more numbers (eight children, 25%). The distribution of children improving by two or more also varied as a function of training condition, with seven of the eight children who improved by two or more numbers being in the counting condition (*Binomial statistic*, $p < .05$).

**The use of verbal counting.** We first asked whether the child's likelihood of verbally counting during the task varied as a function of condition—that is, did children count more after counting was modeled for them in the count training familiarization phase? Overall, 66 children (35.5%) engaged in explicit verbal counting during the Card task. Unsurprisingly, there were significantly more children classified as counters in the counting condition than in the control condition, $\chi^2(1, N = 186) = 63.5$, $p < .001$, *Cramer's V* $= .584$. While only 7.5% of control children

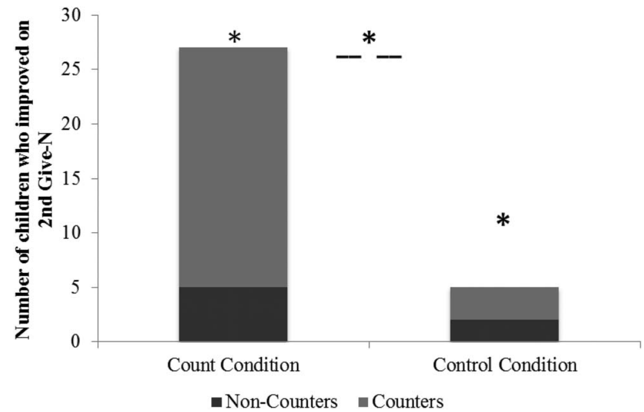


*Figure 2.* In Experiment 1, a significantly greater number of children improved on the second Give-N task after participating in the counting condition compared with the control condition. However, children's counting behavior accounted for this success, such that those children that improved on the follow-up Give-N task were much more likely to be counters (to have engaged in overt counting behavior)—regardless of condition—than noncounters.

($N = 7$) engaged in explicit, verbal counting during the Card task, 63.4% ($N = 59$) of children in the counting condition counted (*Binomial statistic*, $p < .001$). Again, this trend was found across each knower level ($p$'s $< .001$; see Table 4).[8]

Lastly, we asked whether it was the child's counting behavior that mediated the improvements observed in our counting condition. Further regression analyses confirmed that it was counting behavior, and not necessarily the training received in the counting condition, that drove the better performance on both known- and unknown-trials (see Figure 3). When knower-level, counting behavior, and condition were all entered as predictors in two different

[6] Secondary analyses revealed that condition differences were primarily driven by the improved performance of five-knowers in our counting condition. A Condition (Counting vs. Control) × Knower-Level (two-, three-, four-knowers) univariate ANOVA (excluding five-knowers) on percent correct for unknown-number trials revealed a main effect of knower-level, $F(2, 138) = 5.3$, $p = .006$, $\eta_p^2 = .072$, but no main effect condition, $F(1, 138) = .359$, $p > .5$. There was a near-significant interaction between knower-level and condition, $F(2, 138) = 2.9$, $p = .061$, $\eta_p^2 = .04$, accounting for the reversal in performance from two-knowers (performing slightly better in the control condition than the counting condition) to four-knowers (who performed slightly better in the counting condition compared with the control condition).

[7] Notably, the same pattern held when five-knowers were excluded for the analysis, $\chi^2(1, N = 144) = 13.8$, $p < .001$, *Cramer's V* $= .309$.

[8] Post hoc, a random subset of 52 videos (26 in the counting condition and 26 in the control condition, with an equal number of children at each knower-level in each condition) were coded for counting during the first and second Give-N task as well. In line with other reports, we found overall counting rates during this task to be fairly low, not differing significantly from those reported by Sarnecka and Carey (2008; 16%, $p > .5$). Moreover, there were no significant differences in the rate of counting in the first Give-N task as a function of condition (control condition: $M = 7.6\%$ vs. counting condition: $M = 15.4\%$; $p > .05$). However, in line with our findings, we did find a significantly greater number of children in the counting condition engaged in counting during the second Give-N task compared with the control condition (control condition: $M = 3.9\%$ vs. counting condition: $M = 23.1\%$; $p < .05$). Interrater reliability for the coding of counting behavior during both Give-N tasks was 96.2% overall.

Table 3
*Percentage of Participants in Each Knower-Level and Condition who Demonstrated Improved Performance on the Second Give-N Task in Experiment 1*

| | Control condition | Counting condition |
|---|---|---|
| 2-knowers: | 0% | 19.2%* |
| 3-knowers: | 0% | 13.6%^ |
| 4-knowers: | 4.8% | 29.2%* |
| 5-knowers: | 19% | 38.1%* |

^ $p < .1$. * $p < .05$.



regressions predicting known-number trial and unknown-number trial performance, knower-level and counting behavior significantly predicted performance on both known-number and unknown-number trials. In contrast, condition was not a significant predictor in either case ($p$'s $> .2$; Table 5). Formal mediation tests (each controlling for knower-level) confirmed that counting behavior fully mediated the relationship between condition and performance on known number trials (Sobel test $z = 2.31$, $p = .021$) and the relationship between condition and performance on unknown number trials (Sobel test $z = 3.14$, $p = .0017$). Thus, our training manipulation was effective only inasmuch as it encouraged children to count, and subsequently, engaging in counting led to improved performance on both known-number and unknown-number trials (see Figure 3).[9]

Lastly, we examined whether counting behavior on the Card task predicted improvement on the second Give-N task. A chi-square examining the proportion of counters versus non-counters who improved on the second Give-N task revealed that, of those children who improved, the majority were classified as counters (25 of 32, 84.4%); $\chi^2(1, N = 186) = 30.7$, $p < .001$, *Cramer's V* = .406 (see Figure 2). Regression analyses revealed that when all three variables (knower-level, condition, and counting behavior) were entered as predictors, both knower-level and counting behavior (see Table 5) significantly predicted improvement on the second Give-N task, while the child's experimental condition was no longer significant. Again, a formal mediation test (controlling for knower-level) revealed that counting behavior mediated the relationship between condition and improvement on Give-N (Sobel test, $z = 2.79$, $p = .005$).

## Discussion

Encouraging children to count and label sets led to better performance in both the Card task and the second Give-N task. Children in the counting condition across all knower-levels demonstrated improved performance on known-number trials of the Card task relative to those in the control condition, indicating that children benefited from our brief instruction on those trials involving set sizes within their range of mastery. On the other hand, only children classified within the highest knower-levels demonstrated better performance on unknown trials relative to controls. These findings are likely attributed to the fact we asked children to identify a card containing a set size (six) well outside the range of mastery for less-advanced counters (i.e., two-knowers and three-knowers), such that their performance could only be improved by counting inasmuch as they were capable of counting the set sizes presented on these unknown trials. That is, whereas identifying a set size of six relative to other known set sizes was within two- and three-knowers' Zone of Proximal Development (Vygotsky, 1978; i.e., within the realm of abilities that they could accomplish), identifying a set of six relative to significantly larger sets, far outside their range of mastery, was not.

Most importantly, however, is the finding that all four groups of subset knowers were more likely to improve on the second Give-N task after participating in the counting condition, not just the four- and five-knowers (see Table 1). Together, findings suggest that just 5 min of counting instruction for preschoolers can result in improvements in performance both on the specific task that they are trained on, but also on a secondary cardinality task with distinctly different format and task demands.

Furthermore, results of our secondary analyses reveal that the *mechanism* for these dramatic improvements was the direct impact that training had on the children's counting behaviors. That is, children in the counting condition, across all knower-levels, were significantly more likely to engage in explicit verbal counting, and it was these counters who out-performed their noncounting peers in both the Card task and the second Give-N task. In contrast, children in the control condition rarely, if ever, counted. Thus, it appears that the impacts of our counting training manipulation were mediated by children's engagement in verbal counting.

It should be noted that children in our counting condition were exposed to both labeling the cardinality of the set *and* counting, whereas children in our control children were only exposed to labeling of the cardinality of the set. Previously, Mix et al. (2012) found that exposing young children to temporally contiguous labeling *and* counting led to significant improvements in number knowledge above those children exposed to only counting or those exposed to only labeling. In line with their findings, our results reveal that only 5 min of labeling *and* counting sets can lead to improved performance on two distinct cardinality tasks.

## Experiment 2

Modeling cardinal labeling and counting promoted counting behavior that led to improvements in performance across known-number trials, unknown-number trials, and in a Give-N task. Across all four knower-levels, counting instruction increased the amount of overt counting exhibited by participants, and this counting behavior led to improved performance on the Card task.

Performance on unknown-number trials, however, seemed only to improve for those participants with the highest number knowledge. In fact, additional analyses performed excluding five-knowers did not reveal a main effect of condition, suggesting that counting instruction did not significantly alter performance on unknown number trials. One possible explanation for this discrepancy is that children with less number knowledge—who are necessarily less proficient counters—were overwhelmed with having to identify such a large set size outside of their current range of mastery. It is clear that our Card task was challenging, such that

[9] Again, we reran these analyses excluding five-knowers. Regression analyses indicate that performance on both known-number trials ($R^2 = .081$, $p = .008$) and unknown number trials ($R^2 = .118$, $p = .001$) was significantly predicted by counting behavior (known: *Beta* = .224, $p = .031$; unknown: *Beta* = .317, $p = .002$) but not condition ($p$'s $> .1$). Knower-level also predicted performance on unknown (*Beta* = .194, $p = .017$), but not known ($p > .4$) number trials.

Table 4
*Distribution of Counters Within Each Knower-Level and Condition in Experiment 1*

| | Control condition | Counting condition |
|---|---|---|
| 2-knowers: | 0% ($N$ = 0) | 57.7% ($N$ = 15)*** |
| 3-knowers: | 4.0% ($N$ = 1) | 50.0% ($N$ = 11)*** |
| 4-knowers: | 9.1% ($N$ = 2) | 75.0% ($N$ = 18)*** |
| 5-knowers: | 19.0% ($N$ = 4) | 71.4% ($N$ = 15)*** |

*** $p < .001$.

even four- and five-knowers in the control condition performed at chance level. By encouraging two- and three-knowers to engage in a (presumably) challenging new skill of counting in this already cognitively demanding numerical task, they may have become confused and/or felt overwhelmed, especially on unknown-number trials.

In Experiment 2, we explored this possibility by placing two-knowers in a less-challenging Card task, in which they were required to find the card containing three items (similar to Huang et al., 2010) to determine whether performance on less-difficult unknown-number trials would benefit from counting encouragement. Again, children participated in either a counting or non-counting control condition.

Furthermore, to provide a direct comparison between the present findings with those of Huang et al. (2010), we included a second control condition in which the experimenter named the specific animals pictured in each array on the cards (instead of referring to the array collectively as "animals" as in Experiment 1) when prompting children to find the card containing three items (e.g., "Can you find the card with three *dogs*?"). This third condition was included to directly address Huang et al.'s (2010) claims that two-knowers treat new number words like adjectives that are inherently tied to the specific objects they are describing, thereby failing to understand that number words represent abstract concepts, independent of the specific objects depicted in the array. For example, although children may be able to learn to identify a set containing "three *dogs,*" they should be unable to generalize this new knowledge to sets containing "three *cats.*" This second control condition was included to provide a point of comparison between our data and those of Huang et al. (2010), allowing us to explore whether referring to all arrays as "animals" aided performance relative to referring to the specific animals in the array. Thus, we included a *general control* condition (identical to that of Experiment 1) and a *specific control* condition (modeled exactly after Huang et al., 2010) to assess the impact of labeling the type of animals presented in the arrays.

## Method

**Participants.** Again, children were prescreened using the Give-N task to determine their knower-level and only those identified as two-knowers were included in this study. Seventy-two 2- to 5-year-old children identified as two-knowers through the Give-N task participated in this study ($M$ = 33.0 months, $SD$ = 5.88 months; Table 6). Eleven children were excluded due to their failure to complete all test trials (seven), for parental interference (one), or for scoring <50% accuracy on known-number trials (three). Children were again recruited for participation at local museums or for a single visit to our laboratory. After children participated in the Give-N task and were classified as two-knowers, they were then randomly assigned to one of three conditions: (a) counting condition; (b) general control condition (referring to arrays as "animals"); or (c) specific control condition (referring to arrays by specific animal names (e.g., "dogs") as per Huang et al., 2010.

**Materials.**

***Give-N task.*** The materials were identical to that of Experiment 1.

***Card task.*** Stimuli were similar to that of Experiment 1 except that in training, the arrays contained one card with three animals and the other card depicted arrays of one, two, four, six, eight, or 12 animals. Similarly, in test, children saw cards depicting three animals paired with cards depicting one, two, four, six, nine, 12, and 18 animals. Again, the species of animals depicted on the cards varied across trials, such that each trial used a different animal (chickens, pigs, dogs . . .), but within a card pair used for a single trial, all of the animals depicted were identical (e.g., all dogs).

**Procedure.** The procedure of Experiment 2 was identical to that of Experiment 1, with children participating in three tasks: (a) Give-N task; (b) Card task; (c) Give-N task again. The procedure of the first and second Give-N tasks did not differ from that of Experiment 1. The procedure of Card task differed as follows:

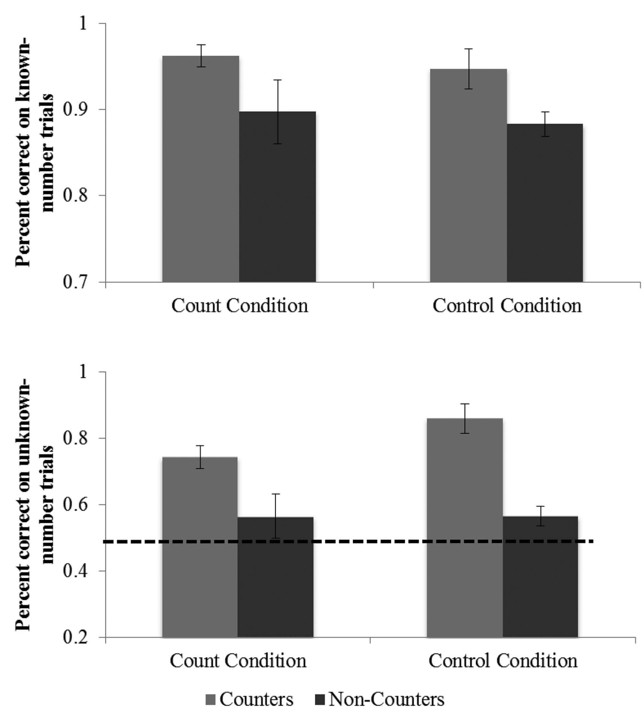


*Figure 3.* In Experiment 1, children who engaged in verbal counting (counters) during the Card task out-performed their noncounting peers (noncounters) in both the control and counting conditions on both known-number trials (top) and unknown number trials (bottom), indicating that explicit counting behavior proved helpful to all children, regardless of condition. The black dashed line indicates chance-level (50%) performance. Error bars reflect standard error of the mean.

Table 5
*Regression Analyses With Knower-Level, Counting Behavior, and Condition Entered as Predictors of Success on Known- and Unknown-Number Trials on the Card Task and on Improvement on the Second Give-N Task*

| Model 1: Knower-level and condition entered as predictors | | |
|---|---|---|
| Outcome variable | Knower-level | Condition |
| Known-number trials | β = .171, $p$ = .017 | β = .216, $p$ = .003 |
| Unknown-number trials | β = .364, $p$ < .001 | β = .159, $p$ = .02 |
| Improvement on Give-N | β = .272, $p$ < .001 | β = .312, $p$ < .001 |

| Model 2: Knower-level, condition, and counting behavior entered as predictors | | | |
|---|---|---|---|
| Outcome variable | Knower-level | Condition | Counting behavior |
| Known-number trials | β = .139, $p$ = .053 | β = .092, $p$ > .2 | β = .213, $p$ = .016 |
| Unknown-number trials | β = .323, $p$ < .001 | β = .002, $p$ > .9 | β = .276, $p$ = .001 |
| Improvement on Give-N | β = .229, $p$ = .001 | β = .144, $p$ = .076 | β = .287, $p$ = .001 |



***Card task: Familiarization for counting condition.*** The procedure was identical to that of the counting condition of Experiment 1, except that children were always asked to identify the card containing three animals using counting (e.g., "This card has *three* animals on it! See, one, two, three! But this [other] card does not have *three* animals on it. See, one, two, three, four . . ."). Again, the experimenter did not refer to the type of animals when counting arrays of card pairs and children did not move to the test phase of the Card task until they had actively participated by counting the cards correctly or selecting the correct target card from each pair.

***Familiarization for general control condition.*** Again, the procedure was identical to that of the control condition of Experiment 1, except that children were asked to identify the set containing three animals, with no extra encouragement to count (e.g., "This card has three animals on it! But this [other] card does not have three animals on it!"]. The experimenter did not refer to the types of animals when counting arrays and children did not move to the test phase of the Card task until they had participated actively by selecting the target card from each pair.

***Familiarization for specific control condition.*** The procedure of this third condition was identical to the procedure of the general control condition except the experimenter referred to the types of animal presented on each card pair when identifying the target number presented (e.g., "This card has three chickens! But this [other] card does not have three chickens!").

***Card task: Test.*** The procedure was identical to that of Experiment 1, except that children were asked to find the card with three items on it (e.g., "Which card has three animals on it?"). As in familiarization, in the counting and general control conditions, the experimenter referred to items on the cards as "animals" throughout test; in contrast, in the specific control condition, the experimenter referred to the specific species of animal depicted on cards ("Which card has three cows on it?").

Table 6
*Distribution of Participants Across Conditions in Experiment 2*

| | Counting | General control | Specific control |
|---|---|---|---|
| *N* | 25 | 25 | 22 |
| *Mean* age in months (*SD*) | 33.4 (5.9) | 33.3 (7.0) | 32.3 (4.5) |

**Data coding and analyses.** As in Experiment 1, children who spontaneously attempted to engage in explicit, verbal counting were coded as "counters" and children who did not engage in verbal counting were coded as "noncounters." Two independent observers coded children's counting strategies for a subset of the participants (20%) and intercoder reliability was 94.1%. As in Experiment 1, any disagreements were settled by a third-party rater and children's counting behavior was classified based on the agreement of two of the three raters.

Again, performance on the Card task was divided into performance on known-number trials (three vs. one and two) and unknown-numbers trials (three vs. four, six, nine, 12, 18).

## Results

### Performance on the Card task.

***Known-number trials.*** Children in all three conditions performed significantly above chance level (50%) on known-number trials ($M_{counting}$ = 92.0%; $M_{general}$ = 79.0%; $M_{specific}$ = 75.0%; all $p$'s < .001, all Cohen's $d'$s > 2.2). A one-way ANOVA investigating the effects of Condition on percent correct for known-number trials revealed a main effect of condition, $F(2, 69) = 4.68$, $p = .012$, $\eta_p^2 = .119$. In line with findings from Experiment 1, children in the counting condition performed significantly better than those in the two control conditions (LSD post hoc tests $p$'s < .03, *Cohen's* $d'$s > .6; Figure 4). Performance on known-number trials in the general control and specific control conditions did not differ ($p > .4$, *Cohen's* $d = .18$).

***Unknown-number trials.*** A one-way ANOVA exploring the effects of condition on percent correct for unknown-number trials revealed no main effect of condition, $F(2, 69) = .964$, $p > .3$, $\eta_p^2 =$ .027. Although not significant, the direction of findings was consistent with predictions, with children in the counting condition ($M$ = 68.0%) performing more accurately than children in the general control condition ($M$ = 66.4%), who in turn performed better than children in the specific control condition ($M$ = 58.6%; Figure 4). Further analyses revealed that whereas children in both the counting and general control conditions performed signifi-

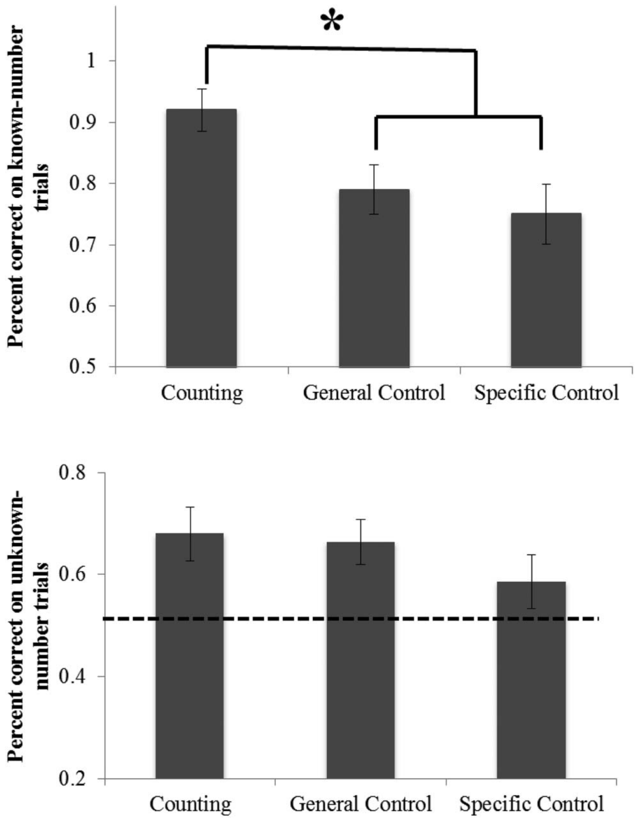


*Figure 4.* Results from Experiment 2 indicate that two-knowers again benefited from our count training procedure on known-number trials (top), but not on the more difficult unknown-number trials (bottom). The black dashed line indicates chance-level (50%) performance. Error bars reflect standard error of the mean.

cantly above chance levels (50%; $p$'s $< .003$, *Cohen's d*'s $> 1.3$), children in the specific control condition did not ($p > .1$, replicating findings of Huang et al., 2010).

**Second Give-N performance.** Analyses revealed improvements in performance on the second Give-N task as a function of condition, $\chi^2(2, N = 72) = 7.07$, $p = .029$, *Cramer's V* = .313. Importantly, more children in the counting condition ($N = 5$, 20.0%) improved on the second Give-N task, compared (marginally) with the general control ($N = 1$, 4.0%); $\chi^2(1, N = 50) = 3.14$, $p = .076$, *Cramer's V* = .246, and specific control ($N = 0$); $\chi^2(1, N = 50) = 4.37$, $p = .037$, *Cramer's V* = .324.

**The use of verbal counting.** Again, the number of children who chose to engage in verbal counting varied as a function of condition, $\chi^2(2, N = 72) = 24.1$, $p < .001$, *Cramer's V* = .578. There were significantly more children who were classified as counters in the counting condition ($N = 18$, 72.0%) than in the general control ($N = 5$, 20.0%); $\chi^2(1, N = 50) = 13.6$, $p < .001$, *Cramer's V* = .522, or the specific control ($N = 2$, 8.0%); $\chi^2(1, N = 47) = 18.9$, $p < .001$, *Cramer's V* = .635, conditions. In contrast to Experiment 1, counting behavior did not appear to drive the improvement in performance in the counting condition on known-number trials. Univariate ANOVAs exploring the effects of both condition and counting behavior on accuracy on known-number trials and on unknown-number trials revealed no main effects or interactions involving the counting variable ($p$'s $> .1$), A main effect of condition in the known-number trial analysis, $F(2, 66) = 6.32$, $p = .003$, $\eta_p^2 = .161$, suggested that it was the counting instruction—not the increased levels of counting behavior—which led to better performance. Counting behavior also did not appear to significantly impact a child's propensity to improve on the second Give-N task (counting behavior: *Beta* = .089, $t(69) = .656$, $p > .5$, [−.211, −.025]).

**Comparing performance across experiments: Find 3 versus Find 6.** We compared two-knowers' performance in the counting and general control conditions across Experiments 1 and 2 to determine how task difficulty (i.e., Find 3 [Experiment 2] vs. Find 6 [Experiment 1]) may have contributed to the pattern of results obtained. Two Condition (Counting, General Control) × Experiment (two) ANOVAs were conducted (for known-number trial performance and unknown-number trial performance separately). Results of the known-number trial analysis revealed a main effect of condition, $F(1, 97) = 6.56$, $p = .012$, $\eta_p^2 = .119$, and no other main effects or interactions ($p$'s $> .2$). In contrast, analyses of unknown-number trial performance revealed that it was experiment (i.e., task difficulty), $F(1, 97) = 10.1$, $p = .002$, $\eta_p^2 = .094$; Exp. 1: $M = 53.8\%$; Exp. 2: $M = 67.2\%$, and not condition, $F(1, 97) = .81$, $p$'s $> .2$, $\eta_p^2 = .008$, that predicted children's accuracy in the Card task.

Analyses of improvement on the follow-up Give-N task showed a similar pattern to that of the known-number trials with condition, $F(1, 97) = 8.44$, $p = .005$, $\eta_p^2 = .08$, alone predicting improvement on the second Give-N task (Experiment did not matter $p > .6$). Whereas a greater proportion of two-knowers improved in the counting than the general control condition, a statistically *similar* proportion of two-knowers in the counting conditions of Experiments 1 and 2 ($M = 19.2\%$ vs. $M = 20.7\%$, $p > .9$) improved and a statistically similar proportion improved in the general control conditions of Experiments 1 and 2 ($M = 0\%$ vs. $M = 4.0$, $p > .3$). Thus, improvement on the Give-N task was not necessarily tethered to specific experiences in the Card task, such that, neither experience with counting to higher numbers in the Find 6 condition nor greater success with counting in the Find 3 condition led to a higher frequency of improvement on the Give-N task.

## Discussion

As in Experiment 1, two-knowers in Experiment 2 who participated in the counting condition outperformed their peers from either control condition on known number trials of the Card task and on the second Give-N task. Thus, results of this experiment align with those of Experiment 1 revealing that just 5 min of counting encouragement and practice can lead to improvements in performance in two distinct counting tasks, even for less-proficient counters.

Surprisingly, despite the reduced demands of the Find 3 task of Experiment 2 (compared to the Find 6 task in Experiment 1), counting instruction again did not have any noticeable impact on performance on unknown number trials. One possibility may be that our counting training procedure was not salient enough to drive a change in performance on these trials. However, given that our training manipulation led to significantly better performance on known-number trials and the second Give-N in both experiments, this does not seem likely.

Performance on the second Give-N task again revealed that encouraging children to count makes it more likely that two-knowers will correctly produce a set of three (or more) objects. Surprisingly, despite the fact that the counting instruction was performed in the context of the Card task, results of both experiments indicate that, at least for these young subset-knowers, our counting manipulation significantly impacted performance in a different cardinality task with entirely distinct task demands and stimuli. Thus, two-knowers readily and flexibly apply the count procedures to novel numerical contexts, and doing so results in substantial performance improvements.

Lastly, results of our Card task in the specific control condition replicated those of Huang et al. (2010), with chance-level responding on unknown-number trials for two-knowers in a Find 3 task where specific animals were named. In contrast, children in our counting and general control conditions both performed significantly above chance. Although a main effect of Condition was not obtained on these unknown-number trials, the chance-level responding in the specific control coupled with the above-chance performance in the general control and counting conditions suggest the use of the broad term "animals" may have helped to slightly boost children's performance over that of children who heard specific animal names when asked to find the card containing "three" items. In line with Huang et al.'s (2010) conclusions, findings reveal that two-knowers in the specific control condition failed to map newly learned words in their count sequence to the fully abstract concept of that number and may be more inclined to think of newly acquired number words as adjectives, tied to the specific items they described. Alternatively, it may be the case that hearing a novel animal name used on every trial may prove distracting to children when engaging in numerical tasks (Kaminski & Sloutsky, 2013). If so, it may not be that hearing broad terms such as "animal" boosted performance; in contrast, it may be that hearing novel animal names on every trial may detract from performance. Regardless, our data extend these findings by suggesting that simply referring to arrays abstractly across sets ("animals") may result in slightly improved performance in numerical tasks, possibly by encouraging children to pay more attention to the numerical properties of the sets and not the items within each set.

## General Discussion

Together, results from these two experiments demonstrate that a short (5-min) labeling and counting instruction manipulation can influence children's performance on numerical tasks. Children of all knower-levels benefited from the brief counting instruction on the Card task. Across knower-levels and experiments, children in our counting condition uniformly out-performed children in the control group on known-number trials. Although the pattern of performance on unknown-number trials was less consistent, more proficient counters (four-knowers and five-knowers) appeared to benefit from our training condition on these trials. Children in our counting condition were not only more accurate on the Card task (which involved the same numerical context as that of training), but children of all knower-levels, in both experiments, were also significantly more likely to improve on a second Give-N task—a markedly different task involving distinct stimuli and task demands—compared with controls. Together, evidence suggests that when encouraged to count, children are more likely to both identify and produce set sizes outside of their range of mastery. Despite claims suggesting the count acquisition process is a rigid, stage-like process (e.g., Le Corre & Carey, 2007), our data indicate that children's early counting abilities can be easily impacted by short instruction, suggesting that perhaps the count acquisition process may be more malleable than previously thought.

Secondary analyses confirmed that the mechanism behind children's improvements were changes in the child's likelihood of engaging in counting behavior. When counting and labeling were modeled, children were significantly more likely to engage in overt counting, even after the experimenter stopped modeling the behavior for them. It appears as though our counting instruction reminded children of the strategy necessary to succeed in identifying numerical sets, specifically those outside of their range of mastery (i.e., unknown-number trials). Moreover, in Experiment 1, counting behavior fully mediated improvements in performance on all three dependent measures (known-number trials, unknown number trials, and the second Give-N task). Of the children who improved on the second Give-N task, the majority of those children engaged in explicit verbal counting during the Card task, after being placed in the counting condition. Thus, children who counted were not only more successful in cardinality identification in the Card task, but they generalized this skill to a different task requiring the production of numerical sets.

Overall, the fact that children were more accurate on the same task in which they were trained to succeed may not be incredibly surprising—they were essentially taught to the test. What is remarkable, however, is that our brief counting instruction also led to improvements in a distinctly different task (Give-N), which was markedly different from the Card task in terms of task demands (e.g., set comparison vs. set production), nature of the items counted (two-dimensional images vs. three-dimensional toys), number of items to be counted (sets outside of the child's mastery vs. within), and the level of child engagement (single choice response vs. jumping animals into a pond to create a set).[10] Because the nature of the Card task and Give-N task differed so substantially, it is remarkable that the benefits of our training generalized across tasks. As such, these findings provide further support to suggest that children were not simply mimicking a meaningless procedure, but instead that they already had some sense of how counting works.

We believe that the present findings speak more broadly in support of the Continuity Hypothesis. Had children simply been parroting the counting behavior without any conceptual understanding (as discontinuity theorists might posit), it is unlikely that performance in the Give-N task would have benefited substantially from the training procedure. That is, subset-knowers should not benefit from counting instruction, as their status would indicate a lack of conceptual understanding that should remain unchanged after five minutes of counting practice and instruction. In contrast,

[10] Of note, the Card task and the Give-N tasks both used animals as their stimuli, possibly providing some parallels for children in our studies. Future work should include other measures of transfer (e.g., tasks that do not include animals) across categories of objects to assess the extent of children's generalization of the counting training.

a continuity theorist might posit that because subset-knowers are primarily limited in their procedural understanding of counting, a reminder and practice with this skill should be beneficial. We find this to be the case, with children succeeding in both the Card task and Give-N task following count instruction. In line with the Continuity Hypothesis, children must have had some basic (albeit raw) conceptual understanding of how and why counting works in order to know when and how to generalize this behavior to this novel task. Consistent with Vygotsky's Zone of Proximal Development (Vygotsky, 1978), our results suggest that rather than teaching children an entirely new concept—for example, an entirely new number outside the range of mastery or a new skill such as counting—the present training paradigm simply drew a connection between performing a numerical task and engaging in counting behavior (something of which the child likely had some basic conceptual understanding). Although children had an understanding of some number words (per their knower-level) and also likely had a basic understanding of the counting procedure, they did not naturally pair these two foundational concepts on their own (as seen in the control condition, very few children engaged in verbal counting). In contrast, children in the counting condition not only used this short modeling session to engage in verbal counting themselves, but it was this engagement that specifically led to bolstered performance on unknown-number trials and improvement on the second Give-N task.

Critically, this short, 5-min counting instruction and practice did not promote all children to a CP-knower status. In fact, we find that most children improved by only one or two numbers on the secondary Give-N task. This may seem surprising, given that continuity theorists suggest that teaching or reminding children to use counting should promote their procedural knowledge, thereby allowing them to perform better following such an intervention. Why did children in the present study not progress past just a few set sizes in the secondary Give-N task? Why did children across knower-levels not easily give sets containing six items, as we might expect? It is likely that children's counting skills (before the intervention) differed across knower levels. That is, two- and three-knowers are generally not able to count out loud as high, or as accurately, as four- and five-knowers. Children with higher number knowledge may also be better versed at counting and thus may be more likely implement it correctly. Moreover, the time-frame of the counting instruction and practice may simply have not been sufficient to result in dramatic increases in performance on the secondary Give-N task. The procedure used in this task took less than 10 min. In fact, after *6 weeks* of counting training, Mix et al. (2012) only observed *small* increments of progress in their follow-up Give-N task, with children only able to produce slightly larger set sizes than previously demonstrated. More work is needed to examine whether qualitatively or quantitatively increased input would lead to more children proficiency giving sets of six (or more) on a secondary Give-N task.

The Give-N task has been broadly accepted as a standard in the field used for assessing a child's level of counting proficiency (Le Corre & Carey, 2007; Wynn, 1990, 1992). Improvement on this task reflects an increase in the child's knower-level classification. Thus, improvements observed over the course of our short experimental session could be explained in two ways: (a) children acquired new number knowledge during our training procedure or (b) children generally underperform on the Give-N task when not encouraged to count. While it is possible that our brief instruction encouraged children to form a deeper connection between the number words the experimenter used to those in their count list, we think this is unlikely the root cause of these improvements given the brief practice that they received. Instead, it is more probable that our training procedure served to remind children of the tool necessary for success in these numerical tasks—that is, we reminded them to count. If so, children who may have been on the verge of advancing to a subsequent knower-level may have been poised to benefit the most from our counting condition. Regardless, if performance on the Give-N task can be influenced by simply encouraging the children to count, it seems that a child's propensity to count when this is highlighted for them may be an important variable to consider when evaluating children's counting abilities. This additional individual difference variable could potentially be an important predictor of the speed with which children acquire new number words and/or later mathematical outcomes.

Importantly, our results suggest that there should be greater attention paid to the counting behavior of children. Previous work indicates that mathematical and quantitative language represents an important link to early quantitative representations (Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006; Miller & Gelman, 1983). In fact, Klibanoff et al. (2006) highlight the importance of "math talk" by teachers of early counters, suggesting that the amount and type of input given to children about counting and cardinality heavily impacts their understanding of and ability to enumerate large set sizes. Our findings build on this work with important educational implications, such that it is possible that a few minutes of counting practice or modeling by a teacher prior to math or number lessons may promote a focus on counting when engaging in these number lessons. In this vein, however, future research should examine whether significantly shorter training paradigms—such as the one used in the present study—similarly have the potential to impact long-term numerical understanding, or whether more substantial training is needed to find such effects. For example, if children were tested an hour later, the following day, or even the following week, would any gains be maintained? Relatedly, future studies should assess the degree of transfer.

Several studies have identified an important link between early counting proficiency and later math achievement (Duncan et al., 2007; Geary, 2011; Stock et al., 2009). By studying how and when children choose to count, we can expand our knowledge of how children transition from an exclusively preverbal understanding of number to a verbal one which, in turn, can aid in the development of methods to target children with early math difficulties, or simply to give children a general head start in the classroom.

## References

Bar-David, E., Compton, E., Drennan, L., Finder, B., Grogan, K., & Leonard, J. (2009). Nonverbal number knowledge in preschool-age children. *Mind Matters: The Wesleyan Journal of Psychology, 4,* 51–64.

Carey, S. (2001). Evolutionary and ontogenetic foundations of arithmetic. *Mind & Language, 16,* 37–55. http://dx.doi.org/10.1111/1468-0017.00155

Carey, S. (2009). *The origin of concepts*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195367638.001.0001

Condry, K. F., & Spelke, E. S. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental*

*Psychology: General, 137,* 22–38. http://dx.doi.org/10.1037/0096-3445.137.1.22

Cordes, S., & Gelman, R. (2005). The young numerical mind: When does it count? In J. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 127–142). London, UK: Psychology Press.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43,* 1428–1446. http://dx.doi.org/10.1037/0012-1649.43.6.1428

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8,* 307–314. http://dx.doi.org/10.1016/j.tics.2004.05.002

Fluck, M., & Henderson, L. (1996). Counting and cardinality in English nursery pupils. *The British Journal of Educational Psychology, 66,* 501–517. http://dx.doi.org/10.1111/j.2044-8279.1996.tb01215.x

Fuson, K. C., Secada, W. G., & Hall, J. W. (1983). Matching, counting, and conservation of numerical equivalence. *Child Development, 54,* 91–97. http://dx.doi.org/10.2307/1129865

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43–74. http://dx.doi.org/10.1016/0010-0277(92)90050-R

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47,* 1539–1552. http://dx.doi.org/10.1037/a0025510

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

Gelman, R., & Tucker, M. F. (1975). Further investigations of the young child's conception of number. *Child Development, 46,* 167–175. http://dx.doi.org/10.2307/1128845

Gunderson, E. A., Spaepen, E., Gibson, D., Goldin-Meadow, S., & Levine, S. C. (2015). Gesture as a window onto children's number knowledge. *Cognition, 144,* 14–28. http://dx.doi.org/10.1016/j.cognition.2015.07.008

Gunderson, E. A., Spaepen, E., & Levine, S. C. (2015). Approximate number word knowledge before the cardinal principle. *Journal of Experimental Child Psychology, 130,* 35–55. http://dx.doi.org/10.1016/j.jecp.2014.09.008

Huang, Y. T., Spelke, E., & Snedeker, J. (2010). When is *four* far more than *three*? Children's generalization of newly acquired number words. *Psychological Science, 21,* 600–606. http://dx.doi.org/10.1177/0956797610363552

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45,* 337–374. http://dx.doi.org/10.1016/S0010-0285(02)00500-5

Kaminski, J. A. (July 2015). *Young children's understanding of the successor function*. Proceedings of the XXXVII Annual Conference of the Cognitive Science Society Meeting, Pasadena, CA: Cognitive Science Society.

Kaminski, J. A., & Sloutsky, V. M. (2013). Extraneous perceptual information interferes with children's acquisition of mathematical knowledge. *Journal of Educational Psychology, 105,* 351–363. http://dx.doi.org/10.1037/a0031040

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk." *Developmental Psychology, 42,* 59–69. http://dx.doi.org/10.1037/0012-1649.42.1.59

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition, 105,* 395–438. http://dx.doi.org/10.1016/j.cognition.2006.10.005

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,* 130–169. http://dx.doi.org/10.1016/j.cogpsych.2005.07.002

Lee, M. D., & Sarnecka, B. W. (2010). A model of knower-level behavior in number concept development. *Cognitive Science, 34,* 51–67. http://dx.doi.org/10.1111/j.1551-6709.2009.01063.x

Lipton, J. S., & Spelke, E. S. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child Development, 76,* 978–988. http://dx.doi.org/10.1111/j.1467-8624.2005.00891.x

Lipton, J. S., & Spelke, E. S. (2006). Preschool children master the logic of number word meanings. *Cognition, 98,* B57–B66. http://dx.doi.org/10.1016/j.cognition.2004.09.013

Miller, K., & Gelman, R. (1983). The child's representation of number: A multidimensional scaling analysis. *Child Development, 54,* 1470–1479.

Miller, K. F., & Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Cognitive Development, 2,* 279–305. http://dx.doi.org/10.1016/S0885-2014(87)90091-8

Mix, K. S., Sandhofer, C. M., Moore, J. A., & Russell, C. (2012). Acquisition of the cardinal word principle: The role of input. *Early Childhood Research Quarterly, 27,* 274–283. http://dx.doi.org/10.1016/j.ecresq.2011.10.003

Munn, P., & Stephen, C. (1993). Children's understanding of number words. *The British Journal of Educational Psychology, 63,* 521–527. http://dx.doi.org/10.1111/j.2044-8279.1993.tb01077.x

Mussolin, C., Nys, J., Content, A., & Leybaert, J. (2014). Symbolic number abilities predict later approximate number system acuity in preschool children. *PLoS One, 9,* e91839. http://dx.doi.org/10.1371/journal.pone.0091839

Negen, J., & Sarnecka, B. W. (2012). Number-concept acquisition and general vocabulary development. *Child Development, 83,* 2019–2027. http://dx.doi.org/10.1111/j.1467-8624.2012.01815.x

Nikoloska, A. (2009). Development of the cardinality principle in Macedonian preschool children. *Psihologija, 42,* 459–475. http://dx.doi.org/10.2298/PSI0904459N

Opfer, J. E., Thompson, C. A., & Furlong, E. E. (2010). Early development of spatial-numeric associations: Evidence from spatial and quantitative performance of preschoolers. *Developmental Science, 13,* 761–771. http://dx.doi.org/10.1111/j.1467-7687.2009.00934.x

Petersen, L. A., McNeil, N. M., Tollaksen, A. K., Boehm, A. G., Hall, C. J., Carrazza, C., & Devlin, B. L. (2014). Counting practice with pictures, but not objects, improves children's understanding of cardinality. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 2633–2637). Austin, TX: Cognitive Science Society.

Piantadosi, S. T., Jara-Ettinger, J., & Gibson, E. (2014). Children's learning of number words in an indigenous farming-foraging group. *Developmental Science, 17,* 553–563.

Pinhas, M., Donohue, S. E., Woldorff, M. G., & Brannon, E. M. (2014). Electrophysiological evidence for the involvement of the approximate number system in preschoolers' processing of spoken number words. *Journal of Cognitive Neuroscience, 26,* 1891–1904. http://dx.doi.org/10.1162/jocn_a_00631

Posid, T., & Cordes, S. (2014). Verbal counting moderates perceptual biases found in children's cardinality judgments. *Journal of Cognition and Development, 16,* 621–637. http://dx.doi.org/10.1080/15248372.2014.934372

Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition, 108,* 662–674. http://dx.doi.org/10.1016/j.cognition.2008.05.007

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

Sarnecka, B. W., & Lee, M. D. (2009). Levels of number knowledge during early childhood. *Journal of Experimental Child Psychology, 103,* 325–337. http://dx.doi.org/10.1016/j.jecp.2009.02.007

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

Slusser, E. B., & Sarnecka, B. W. (2011). Find the picture of eight turtles: A link between children's counting and their knowledge of number word semantics. *Journal of Experimental Child Psychology, 110,* 38–51. http://dx.doi.org/10.1016/j.jecp.2011.03.006

Stock, P., Desoete, A., & Roeyers, H. (2009). Master of the counting principles in toddlers: A Crucial step in the development of budding arithmetic abilities? *Learning and Individual Differences, 19,* 419–422. http://dx.doi.org/10.1016/j.lindif.2009.03.002

Vygotsky, L. (1978). Interaction between learning and development. In V. Pareto (Ed.), *Mind and society* (pp. 79–91). Cambridge, MA: Harvard University Press.

Wagner, J. B., & Johnson, S. C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition, 119,* 10–22. http://dx.doi.org/10.1016/j.cognition.2010.11.014

Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology, 43,* 217–242. http://dx.doi.org/10.1006/cogp.2001.0764

Wylie, J., Jordan, J. A., & Mulhern, G. (2012). Strategic development in exact calculation: Group and individual differences in four achievement subtypes. *Journal of Experimental Child Psychology, 113,* 112–130. http://dx.doi.org/10.1016/j.jecp.2012.05.005

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36,* 155–193. http://dx.doi.org/10.1016/0010-0277(90)90003-3

Wynn, K. (1992). Children's acquisition of number words and the counting system. *Cognitive Psychology, 24,* 220–251. http://dx.doi.org/10.1016/0010-0285(92)90008-P

Zur, O., & Gelman, R. (2004). Young children can add and subtract by predicting and checking. *Early Childhood Research Quarterly, 19,* 121–137. http://dx.doi.org/10.1016/j.ecresq.2004.01.003

Received July 31, 2015
Revision received September 26, 2017
Accepted October 4, 2017 ■