

## Research Reports

## Parallel Individuation Supports Numerical Comparisons in Preschoolers

Pierina Cheung<sup>\*ab</sup>, Mathieu Le Corre<sup>c</sup>

[a] Department of Psychology, University of Waterloo, Waterloo, Canada. [b] National Institute of Education, Nanyang Technological University, Singapore. [c] Centro de Investigación en Ciencias Cognitivas, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, México.

## Abstract

While the approximate number system (ANS) has been shown to represent relations between numerosities starting in infancy, little is known about whether parallel individuation – a system dedicated to representing objects in small collections – can also be used to represent numerical relations between collections. To test this, we asked preschoolers between the ages of 2 ½ and 4 ½ to compare two arrays of figures that either included exclusively small numerosities (< 4) or exclusively large numerosities (> 4). The ratios of the comparisons were the same in both small and large numerosity conditions. Experiment 1 used a between-subject design, with different groups of preschoolers comparing small and large numerosities, and found that small numerosities are easier to compare than large ones. Experiment 2 replicated this finding with a wider range of set sizes. Experiment 3 further replicated the small-large difference in a within-subject design. We also report tentative evidence that non- and 1-knowers perform better on comparing small numerosities than large numerosities. These results suggest that preschoolers can use parallel individuation to compare numerosities, possibly prior to the onset of number word learning, and thus support previous proposals that there are numerical operations defined over parallel individuation (e.g., Feigenson & Carey, 2003; <https://doi.org/10.1111/1467-7687.00313>).

**Keywords:** parallel individuation, analog magnitudes, numerical comparison, small, large, preschoolers

Journal of Numerical Cognition, 2018, Vol. 4(2), 380–409, doi:10.5964/jnc.v4i2.110

Received: 2016-12-12. Accepted: 2017-09-20. Published (VoR): 2018-09-07.

\*Corresponding author at: National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore, 637616. E-mail: cheung.pierina@gmail.com



This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

According to structuralism, an influential theory of the nature of mathematical objects, the natural numbers are nothing but a system of relations (e.g., Shapiro, 2000). Therefore, accounts of our knowledge of the natural numbers and of its developmental origins cannot be limited to explaining how we represent individual numbers and how we acquire these representations. They must also explain how we represent relations between them and how these representations develop. Randy Gallistel (1989) put the point quite forcefully:

“If the only sense in which the brain represents number is that there is a sensory/perceptual mapping from numerosity to brain states (the activities of detectors for specific numerosities), which make possible simple numerical discriminations, then the brain’s representation of number is a representation in name only. Only if the brain brings combinatorial processes to bear on the neural entities that represent numerosities may we say that the brain represents number in an interesting sense of the term representation.” (p. 159).

One of the possible developmental roots of our representations of relations between numbers is the Approximate Number System (ANS). The ANS represents number approximately – i.e., it does not discriminate all pairs of numerosities with equal precision (Dehaene, 1997; Gallistel & Gelman, 2000; Meck & Church, 1983). Rather, the discrimination precision of the ANS is a function of the ratio of numerosities (e.g., Coubart, Izard, Spelke, Marie, & Streri, 2014; Lipton & Spelke, 2004; Moyer & Landauer, 1967; Xu & Arriaga, 2007; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005). For example, 9-month-olds can discriminate numerosities that differ by a 2:3 ratio (e.g., 6 vs. 9), but not a 3:4 ratio (e.g., 12 vs. 16) (Xu, 2003; Xu & Spelke, 2000; Xu et al., 2005). Starting in infancy, the ANS also represents relations between numerosities such as relative numerosity (Brannon, 2002; Suanda, Tompson, & Brannon, 2008; see Barth et al., 2005, 2006 for evidence from preschoolers), addition and subtraction (McCrink & Wynn, 2004; see also Barth et al., 2005, 2006), and proportions of discrete quantities (McCrink & Wynn, 2007). Although this is controversial, some studies suggest that later numerical reasoning about numerical relations is at least partially rooted in the ANS, even when the reasoning involves manipulating mathematical symbols (e.g., Halberda, Mazzocco, & Feigenson, 2008; Libertus, Feigenson, & Halberda, 2011; Starr, Libertus, & Brannon, 2013a; but see Gilmore et al., 2013; Szűcs, Nobes, Devine, Gabriel, & Gebuis, 2013; for reviews, see De Smedt, Noël, Gilmore, & Ansari, 2013).

However, multiple studies have shown that, in various contexts, infants and adults often do not use the ANS to represent the numerosity of collections when they include 4 or fewer objects. Instead, they use a distinct system whose capacity is limited to collections of up to 3 or 4 objects (Choo & Franconeri, 2014; Feigenson & Carey, 2003, 2005; Feigenson, Carey, & Hauser, 2002; Feigenson, Dehaene, & Spelke, 2004; Lipton & Spelke, 2004; Revkin, Piazza, Izard, Cohen, & Dehaene, 2008; Trick & Pylyshyn, 1994; Xu, 2003; see also Van Herwegen, Ansari, Xu, & Karmiloff-Smith, 2008; Xu, Spelke, & Goddard, 2005). Most agree that this system represents individual objects and that the limit on its capacity comes from the number of objects it can represent in parallel. It is thus often referred to as “parallel individuation.” There is also growing agreement that this system develops early in infancy (e.g., Coubart et al., 2014; Feigenson, Carey, & Hauser, 2002; Hyde & Spelke, 2011).

Some have proposed that, in addition to representing objects, parallel individuation supports numerical operations (Carey, 2009; Feigenson & Carey, 2003, 2005; Le Corre & Carey, 2007, 2008). On this view, representations of objects created by the parallel individuation system can enter into computations of one-to-one correspondence from infancy on (Carey, 2009; Feigenson & Carey, 2003, 2005; see also Uller, Carey, Huntley-Fenner, & Klatt, 1999). For example, to explain how infants can keep track of up to three objects hidden in an opaque box, Feigenson and Carey (2003) suggested that infants represent each hidden object with a unique symbol that is held in working memory, where each symbol functions as a mental tally mark of sorts. Every time they retrieve an object, they match it to an active tally mark in working memory. When each tally mark has been matched to an object, infants stop reaching. Second, infants can use parallel individuation to represent more than one collection at a time. That is, under some conditions, representations of objects can be grouped into up to two or three “chunks” of up to three objects (Moher, Tuerk, & Feigenson, 2012; Zosh, Halberda, & Feigenson, 2011; see Frank & Barner, 2012 for evidence of object-chunking in school-age children).

If parallel individuation does indeed support numerical operations, one might expect that the system could be used to compare distinct collections of objects on the basis of their numerosity. Some studies have shown that infants can use parallel individuation to compare collections. However, these studies found that infants’ comparisons were based on physical attributes of the collections (e.g., their total physical size) rather than on their nu-

merosity (Clearfield & Mix, 1999, 2001; Feigenson, Carey, & Hauser, 2002; Feigenson, Carey, & Spelke, 2002; Xu, et al., 2005). For example, in a quantity choice task, 10- to 12-month-olds were shown small collections of crackers put into two buckets. Feigenson, Carey, and Hauser (2002) found that infants chose the bucket with the larger number of crackers only when there were no more than three in each bucket; when one or both of the buckets contained more than three (e.g., 1 vs. 4, 2 vs. 4, 3 vs. 6), infants chose both buckets equally frequently. The upper limit on infants' performance suggests that parallel individuation was recruited. Nevertheless, it was also found that for comparisons that involved less than four crackers, infants reliably chose the bucket with more cracker stuff rather than more pieces of crackers (e.g., they chose one big cracker with a larger surface area over two pieces of cracker with a smaller combined surface area). Studies requiring that infants retrieve hidden objects from a box suggest that infants can compare a set against another held in working memory (e.g., Feigenson & Carey, 2003), but they do not test whether infants can perform comparisons on two *physically distinct* collections under parallel individuation because they never require infants to do this. Thus, these studies have not shown that parallel individuation can be used to determine which of two collections contains more elements on the basis of number.

To test whether parallel individuation can support numerical comparisons, a study must include numerosities that are small enough to be compared with parallel individuation. But this is not enough, since, in principle, participants can also use the ANS to compare small numerosities (Cordes et al., 2001; Cordes & Brannon, 2009; Starr, Libertus, & Brannon, 2013b). Therefore, the study must also include a way to determine which of the two systems is used to compare the small numerosities. One way to do so is to include pairs of numerosities that can be compared with the ANS but not with parallel individuation, and whose ratio is the same as the ratio of the pairs of small numerosities. Evidence that, despite the fact that all comparisons have the same ratio, performance on comparisons of collections that can be represented with parallel individuation (henceforth, "small numerosities") is significantly different from performance on comparisons of numerosities that can only be represented with the ANS (henceforth, "large numerosities") would suggest that parallel individuation can support numerical comparisons.

Pairs of collections that straddle the boundary between parallel individuation and the ANS must be avoided because the pattern of performance to be expected in such situations is not well known. Unfortunately, the boundary between the two systems is somewhat unclear. Some studies suggest that parallel individuation can hold up to 4 objects, but many others suggest that it cannot hold more than 3 (e.g., Feigenson & Carey, 2005; Feigenson, Carey, & Hauser, 2002). The bulk of the evidence that suggests it can hold up to 4 objects comes from studies of human adults or rhesus monkeys (e.g., Hauser, Carey, & Hauser, 2000; Luck & Vogel, 1997; Pylyshyn & Storm, 1988); only one out of many infant studies suggests that it can hold up to 4 objects (Ross-Sheehy, Oakes, & Luck, 2003). Moreover, a study of non-verbal numerical comparisons in human adults has shown that they use parallel individuation when collections are comprised of 3 or fewer objects, and that they rely on the ANS when they are comprised of 4 or more objects (Choo & Franconeri, 2014). Therefore, we suggest that comparisons should not include collections of 4 objects because the way such collections are represented is unclear. In other words, small numerosities should be limited to collections of up to 3 objects and large numerosities should consist of collections of at least 5 objects.

Previous studies with preschoolers have shown that children can compare two collections on the basis of number, but these studies do not meet the conditions outlined above (Abreu-Mendoza, Soto-Alba, & Arias-Trejo, 2013; Odic, Libertus, Feigenson, & Halberda, 2013; Odic, Pietroski, Hunter, Lidz, & Halberda, 2013; Rousselle

& Noël, 2008; Rousselle, Palmers, & Noël, 2004; Wagner & Johnson, 2011). Some showed clear evidence of comparisons based on numerosity but the pairs of small collections also included collections of more than 3 objects (Abreu-Mendoza, Soto-Alba, & Arias-Trejo, 2013; Cantrell, Kuwubara, & Smith, 2015; Odic, Libertus, et al., 2013; Odic, Pietroski, et al., 2013; Rousselle, Palmers, & Noël, 2004; Wagner & Johnson, 2011). Other studies did include a condition where none of the pairs of small collections contained more than 3 objects but did not include a condition where the two collections in each pair contained at least 5 objects (Barner & Snedeker, 2005, 2006; Brannon & Van de Walle, 2001; Feigenson, 2005; Feigenson, Carey, & Hauser, 2002; Feigenson, Carey, & Spelke, 2002). It is thus not clear whether preschoolers were recruiting the ANS or parallel individuation to compare small collections on the basis of their numerosity in these studies.

To our knowledge, only two studies met all of the demands laid out above. One of these studies tested adults (Choo & Franconeri, 2014). While it provided evidence that parallel individuation can be used to compare numerosities (i.e., at equal ratios, adults were faster at comparing small numerosities than large numerosities), it did not provide evidence that is relevant to the development of this capacity. Cantlon, Safford, and Brannon (2010) tested preschoolers. However, their results are difficult to interpret. Their study tested comparisons with a numerical matching task – i.e., participants were shown a target collection and then had to match it to one of two other collections on the basis of numerosity. The task included only one pair of comparisons with the right design features – namely, a pair comprised of comparisons of small collections and of large collections at the same ratio – 1 vs. 2 and 6 vs. 12. They found that children were better at finding the numerical match for 1 object (out of 1 and 2 objects) than at finding the match for 6 or 12 objects (out of 6 and 12). However, they also found that children were *not* significantly better at finding a match for 2 (out of 1 and 2) than at finding a match for 6 or 12 (out of 6 and 12). Therefore, their results are inconclusive.

The present study aims to contribute to research on the development of parallel individuation as a system for numerical reasoning by applying a numerical comparison task with the right design features to preschoolers, namely two- to four-year-olds. Since, other than one study with inconclusive results (Cantlon et al., 2010), no previous study of numerical comparison in preschoolers had all of the right design features, our study provides the clearest test thus far of whether preschoolers can use parallel individuation to compare numerosities. Moreover, our study included children who were young enough to be at the earliest stages of number word learning, namely children who had not learned the meaning of any of the number words beyond “one.” Thus, it also provides evidence that bears on whether the development of the capacity to use parallel individuation to compare numerosities depends on number word learning.

## Experiment 1

### Method

#### Participants

A total of 99 2½ - to 4½-year-olds participated in this study. Fifty of them were tested on comparisons of small collections only (average age of 3 years 7 months; range: 2 years 7 months – 4 years 7 months; 22 males), and 49 were tested on comparisons of large collections only (average age of 3 years 7 months; range: 2 years 6 months – 4 years 7 months; 30 males). All of the children were recruited in Southwestern Ontario, Canada, and

were predominantly monolingual speakers of English. An additional two children were excluded for always choosing the same side.

### Design and Procedure

Children were tested on a numerical comparison task and on Give-*N*, a standard assessment of number word knowledge (Wynn, 1990). The Give-*N* task was always administered at the end of the testing session.

**Numerical comparison —** The numerical comparison task always started with an experimenter introducing two puppets to the children – a frog and a duck. Then, children were shown a picture of a rectangle and asked to name it, e.g., “Do you know what this is?” If children failed to provide a label, the experimenter suggested one (e.g., block, rectangle), and encouraged children to repeat it. After familiarization with the stimuli, the experimenter showed children two pictures of blocks, placed one in front of each of the two puppets, and said, “Froggie has some [blocks], duckie has some [blocks], who has more [blocks]?” This instruction was repeated for each trial.

The numerical comparison task had two between-subject conditions that differed on the range of numerosities tested: small (< 4) and large (> 4). In each condition, collections that differed by a 1:3 and 2:3 ratio were shown. Children in the small numerosity condition were asked to compare 1 vs. 3 (4 trials) and 2 vs. 3 (4 trials). Children in the large numerosity condition were asked to compare 6 vs. 18 (4 trials) and 6 vs. 9 (4 trials). All collections consisted of red rectangles printed on letter-sized paper. The rectangles in each individual collection were all of the same physical size. However, the relation between the physical size of individual rectangles, their total perimeter and area, and numerosity varied across trials (see Figures 1a, 1b, 1c and 1d). As in previous studies (e.g., Halberda & Feigenson, 2008; Rousselle & Noël, 2008), on half of the trials, the two collections had the same cumulative perimeter and surface area, so that the physical size of the individual rectangles in the collections conflicted with numerosity (i.e., the rectangles in the numerically smaller collection were physically larger). We refer to these trials as “size-number incongruent.” (Figures 1b and 1d). On the other half of the trials, the cumulative perimeter and surface area of the collections was confounded with numerosity (i.e., the collection with a greater number of objects also had a larger cumulative perimeter and a larger surface area; Figures 1a and 1c). We refer to these trials as “size-number congruent.”

In the small numerosity condition, for size-number congruent trials, the cumulative surface area of the collections ranged from 1.5 cm<sup>2</sup> to 12 cm<sup>2</sup>, and their cumulative perimeter ranged from 5.32 cm to 30 cm. For size-number incongruent trials, the cumulative surface area was always 6 cm<sup>2</sup>, and cumulative perimeter was 18 cm. In the large numerosity condition, for size-number congruent trials, the cumulative surface area of the collections ranged from 18 cm<sup>2</sup> to 72 cm<sup>2</sup>, and their cumulative perimeter ranged from 31.92 cm to 144 cm. For size-number incongruent trials, cumulative surface area ranged from 18 cm<sup>2</sup> to 36 cm<sup>2</sup>, and cumulative perimeter ranged from 54 cm to 108 cm.

Pairs of collections of rectangles were presented in one of two item orders. The orderings of correct side (left or right), ratio (1:3 or 2:3), numerosity pair (small condition: 2 vs. 3, 1 vs. 3; large condition: 6 vs. 9, 6 vs. 18), and trial type (size-number congruent or size-number incongruent) were randomized across orders. No two consecutive trials were of the same type or pair of numerosity. No feedback was given.

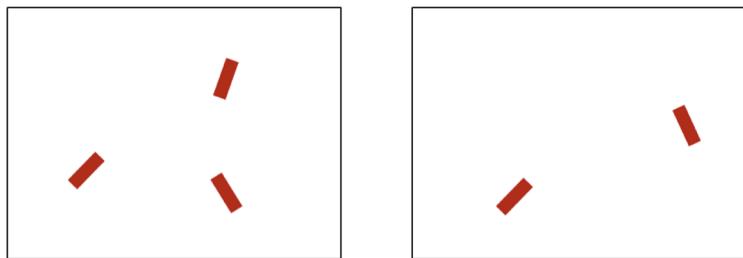


Figure 1a. A size-number congruent trial in the small numerosity condition in Experiment 1.

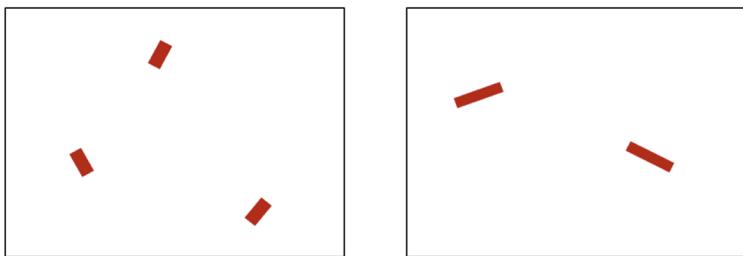


Figure 1b. A size-number incongruent trial in the small numerosity condition in Experiment 1.

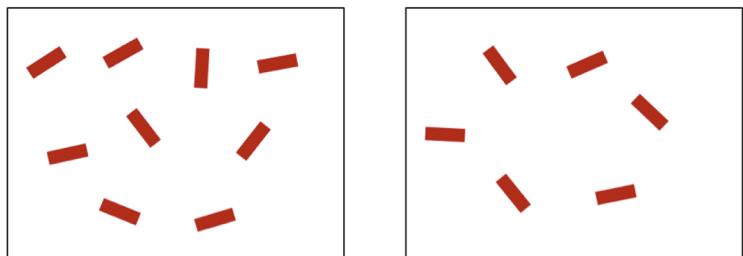


Figure 1c. A size-number congruent trial in the large numerosity condition in Experiment 1.

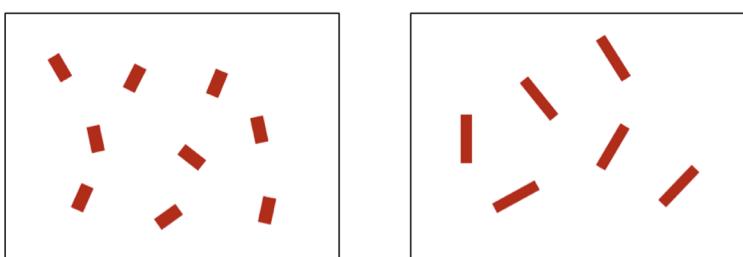


Figure 1d. A size-number incongruent trial in the large numerosity condition in Experiment 1.

**Give-N** — The purpose of this task was to assess children's number word knowledge. Children were first introduced to a puppet, a tub of 10 fish, and a plate. They were then told that the puppet wanted to eat some fish, and the experimenter asked the child to give the puppet  $N$  fish (e.g., "Can you put one fish on the plate?"). After the child gave the puppet some fish, the experimenter asked whether it is  $N$ ; if the child said 'no', s/he was asked to fix it and if the child said 'yes', the experimenter moved onto the next trial. Children were asked to give

1 and then 3 fish on the first two trials. If children succeeded on both, they were asked to give 5 fish. If children failed to correctly give 1 for “one” or 3 for “three”, the experimenter asked for two fish. At this point, if children succeeded in response to a request for  $N$ , the next request was  $N+1$ ; if they incorrectly responded to the request for  $N$ , the next request was for  $N-1$ . The highest numeral requested was “six”.

Children were called ‘ $N$ -knowers’ (e.g., ‘1-knowers’) if they correctly gave  $N$  fish two out of three times when asked for  $N$ , but failed to give the correct number two out of three times for  $N+1$ . Children who failed to give one fish when asked for “one” were classified as ‘non-knowers’. Children who only knew a subset of the number words – i.e., ‘1-knowers’, ‘2-knowers’, ‘3-knowers’, and ‘4-knowers’ – were called ‘subset-knowers’. Children who gave the correct number of fish for all numerals asked for (up to six) were called ‘Cardinal Principle-knowers’, ‘CP-knowers’ for short. In the analyses throughout this paper, we divided children into three knower-level groups: non- and 1-knowers vs. 2- and 3-knowers vs. CP-knowers.

## Results

Across all three experiments, whenever multiple comparisons were performed, the alpha level was adjusted using the Holm-Bonferroni method.

### Give-N

The number of children and the mean age in each knower-level group are presented in Table 1. Age was significantly correlated with children’s knower-level, Pearson’s  $r = .51$ ,  $p < .001$ .

Table 1

*Age and Number of Children in Each Knower-Level Group in Experiment 1*

Knower-Level	Small			Large		
	n	Mean Age	Range	n	Mean Age	Range
Non-knowers	8	3;1	2;9 - 3;8	7	3;5	3;2 - 3;10
1-knowers	12	3;5	2;7 - 4;0	10	3;0	2;6 - 3;7
2-knowers	10	3;5	2;8 - 3;11	12	3;9	3;2 - 4;7
3-knowers	12	3;10	3;1 - 4;7	11	3;8	2;11 - 4;2
CP-knowers	8	3;11	3;2 - 4;7	9	3;11	2;11 - 4;7

### Numerical Comparisons

Our first analysis found no effect of order or gender (all  $ts < -1.26$ ,  $ps > .21$ ), so these variables were not included in subsequent analyses. A logistic mixed effects model was used to analyze the effect of our independent variables on proportion correct responses on the numerical comparison task. We began with a maximal model that included random slopes of Congruence and Ratio by subjects but this model did not converge. We then sequentially removed random effects by Congruence and Ratio and none of the models converged; thus, in our main model, we only included random intercept by subject.

We included 2-way interaction terms involving Number Range (Number Range x Congruence, Number Range x Ratio, Number Range x Knower-Level, and Number Range x Age) because we were primarily interested in the effects of that factor. In particular, the model included Number Range (small vs. large), Size-Number Congruence (congruent vs. incongruent), Ratio (1:3 vs. 2:3), Knower-Level (non- and 1-knowers vs. 2- and 3-know-

ers vs. CP-knowers; CP-knowers as the reference category) and centered Age as fixed effects, with by-subject random intercept. This model did not increase the fit over a model with main effects only,  $\chi^2(5) = 4.82, p = .44$ . A main-effects-only model revealed a main effect of Number Range,  $\beta = -.87, SE = .29, z = -3.01, p = .0026$ . Children were better at comparing small numerosities ( $M = .79, SD = .24$ ) than large numerosities ( $M = .68, SD = .26$ ). There was a main effect of Size-Number Congruence,  $\beta = -.99, SE = .20, z = -5.02, p < .001$ , with better performance on congruent ( $M = .81, SD = .27$ ) than on incongruent trials ( $M = .67, SD = .32$ ). We also found an effect of Age,  $\beta = 1.03, SE = .36, z = 2.89, p = .0039$ , with older children performing better overall. Finally, we found a main effect of Knower-Level, with CP-knowers ( $M = .93, SD = .16$ ) performing better than both non- and 1-knowers ( $M = .58, SD = .25; \beta = -2.26, SE = .57, z = -3.96, p < .001$ ) and 2- and 3-knowers ( $M = .79, SD = .22; \beta = -1.36, SE = .53, z = -2.57, p = .010$ ). Non- and 1-knowers also differed significantly from 2- and 3-knowers ( $t(80) = -4.08, p < .001$ ). No other effects were significant. Figure 2 displays children's performance on small and large comparisons by knower-levels and by size-number congruence.

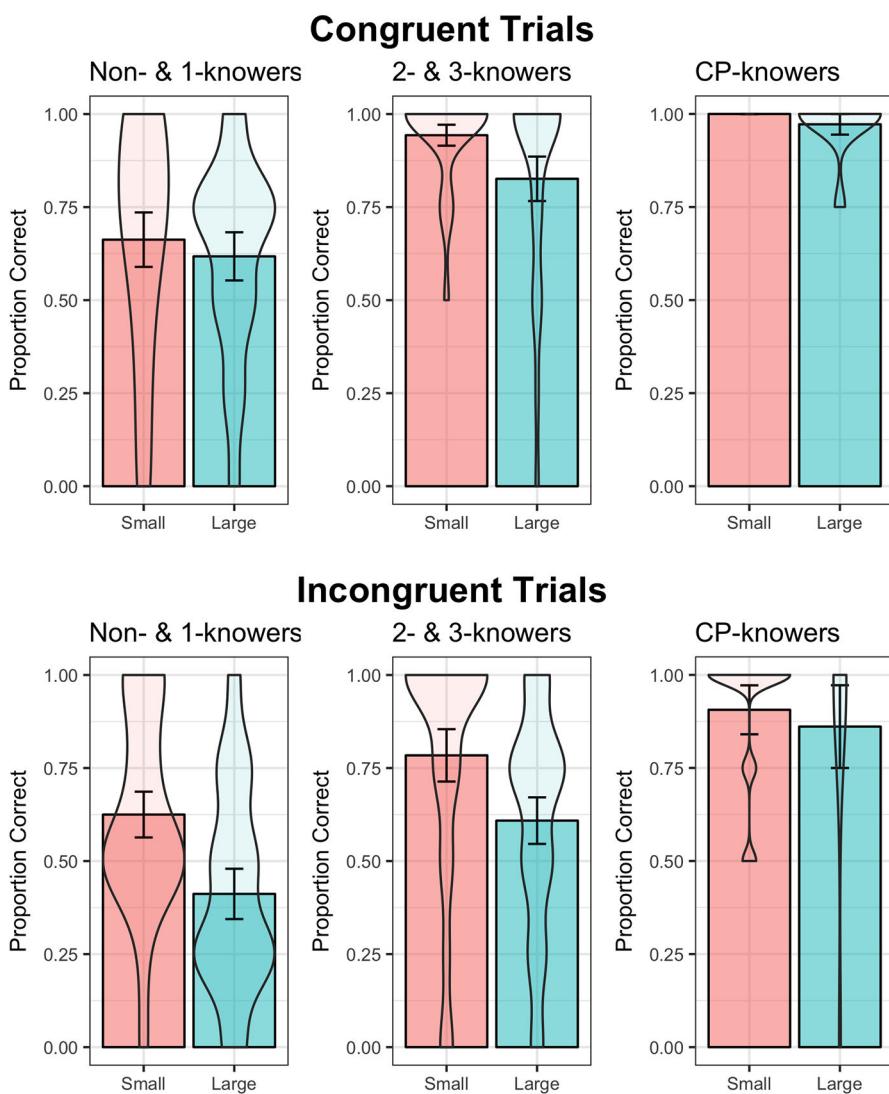


Figure 2. Overall proportion correct on small and large numerosity comparisons on congruent trials (top panel) and incongruent trials (bottom panel), separated by knower-levels in Experiment 1. Violin plots are used to depict the distribution and probability density of the data. Error bars represent standard errors of the mean.

Finally, we asked whether the ability to use parallel individuation to compare numerosities is available even at the earliest stages of number word learning. To test this, we analyzed whether children with minimal number word knowledge – non-knowers and 1-knowers – performed significantly above chance on size-number incongruent trials for small numerosities, and whether their performance for small numerosities was better than that for large numerosities. We combined 2 vs. 3 and 1 vs. 3 comparisons and found that non-knowers and 1-knowers as a group performed significantly above chance on small comparisons,  $M = .63$ ,  $SD = .28$ ,  $t(19) = 2.03$ ,  $p = .028$  (1-tailed).<sup>ii</sup> We also found that they performed significantly better on comparisons of small numerosities (1 vs. 3 and 2 vs. 3) than of large numerosities (6 vs. 18 and 6 vs. 9;  $M = .41$ ,  $SD = .28$ ;  $t(35) = 2.34$ ,  $p = .025$  (2-tailed),  $d = .77$ , 95% CI [0.08, 1.46]).<sup>iii</sup> These results suggest that parallel individuation can support numerical comparisons even in children with little or no knowledge of number word meanings. They also suggest that children's performance on the numerical comparisons task cannot be explained by counting, because these children have not acquired the cardinal principle, and thus cannot have used counting to solve the task (Le Corre, Van de Walle, Brannon, & Carey, 2006; Sarnecka & Carey, 2008; Wynn, 1990, 1992).

## Discussion

Two- to four-year-old children were tested on a non-verbal numerical comparison task in one of two conditions: comparing collections of three or fewer objects or collections of six or more objects. The ratios of the numerosities in both the small and large numerosity comparisons were the same. Despite that, children did not perform equally well on all comparisons. Rather, they were more accurate on small than on large comparisons. This difference also held in children at the earliest stages of number word learning – i.e., non-knowers and 1-knowers. These children performed significantly above chance when they compared small numerosities but not when they compared large ones. Given that the number range effect was observed in children at the earliest stages of number word learning and that these children performed above chance on comparisons of small numerosities even when area and numerosity were not congruent, our results also suggest that the development of the ability to use parallel individuation to compare numerosities does not depend on number word learning.

Two aspects of preschoolers' performance on large numerosity comparisons warrant discussion. First, children at the earliest stages of number word learning (i.e., non- and 1-knowers), and to some extent, 2- and 3-knowers, performed poorly on comparisons of large numerosities (see Figure 2). This seems surprising given that by 9 months, infants can discriminate between large collections that differ by a 2:3 ratio (e.g., Xu & Spelke, 2000). We speculate that the discrepancy between infants and two- to four-year-olds is that the habituation paradigm used with infants cues the fact that object size and summed area are not relevant to the task to a greater extent than our numerical comparison task. In the habituation experiments (e.g., Xu & Spelke, 2000), number stays constant in all habituation arrays but the size of individual objects and their total area vary from array to array. This may signal to infants that numerosity is more relevant to the task than individual object size or total area such that by the time the test arrays are presented, infants can discriminate the habituation arrays from the test arrays on the basis of numerosity. In contrast, aside from the count noun in the question ("Who has more blocks?"), no feature of our design cues children to ignore the size of the objects in the collections on size-number incongruent trials. Thus, although the acuity of children's ANS is sufficiently high to compare numerosities at the ratios we presented, they fail to do so because nothing helps them overcome the interference between object size and numerosity. Negen and Sarnecka (2015) provide evidence that is consistent with this explanation. They presented subset-knowers with numerical comparisons where numerosity conflicted with total area and individual object size. Like us, they found that subset-knowers performed at chance on these comparisons,

despite the fact that the ratios of the comparisons were well within the capacity of young infants. However, they also found that after training subset-knowers to focus on numerosity instead of total area and individual object size, they were able to choose the numerically larger collection instead of the collection with larger individual objects. These findings thus show that their initial failure was due to the interference between object size and numerosity, and that they can overcome this interference when cues (in this case, explicit feedback) are provided.

Second, none of the effects involving ratio were significant. Since this suggests that children were not more accurate on the large comparisons with a large ratio (6 vs. 18) than on the large comparisons with a smaller ratio (6 vs. 9), it may seem to pose problems for our suggestion that children used the ANS to compare large numerosities. That is, one may be concerned that our task was not appropriately designed to test our question of interest because it may not have engaged the ANS. We believe that this concern is unwarranted for two reasons. First, three other studies of non-verbal numerical comparisons in preschoolers with designs similar to ours also failed to find differences of accuracy between ratios similar to ours – i.e., 1:2 and 2:3 (Abreu-Mendoza, Soto-Alba, & Arias-Trejo, 2013; Halberda & Feigenson, 2008; Rousselle & Noël, 2008). Importantly, two of these studies provided positive evidence that their task engaged the ANS – i.e., they found that children were better on comparisons with 1:2 and 2:3 ratios than on comparisons with harder-to-discriminate ratios (Abreu-Mendoza et al., 2013; Halberda & Feigenson, 2008). Moreover, children's average accuracy on the large numerosity comparisons in the present study was similar to that which has been reported in previous studies for children of the same age – i.e., near 70% for comparisons with ratios between 2:3 and 1:2 (Halberda & Feigenson, 2008; Rousselle & Noël, 2008). Therefore, our failure to find a difference between performance on comparisons of 6 vs. 9 (a 2:3 ratio) and of 6 vs. 18 (a 1:3 ratio) does not mean that our task did not engage the ANS. Rather, it is consistent with, perhaps even predicted by, what we know about how ratio affects children's performance when they use the ANS to compare numerosities.

There is no other plausible explanation of children's performance on large comparisons in our study. The only other strategies that are available to adults are counting, and breaking up the large collections into smaller collections of 2 to 4 elements and adding these up. Neither of these strategies can explain why we find the same pattern of results when we restrict our analyses to subset-knowers. Subset-knowers do not know the cardinal principle. Therefore, they cannot have compared the collections by counting them. Moreover, it is highly unlikely that children who do not know the cardinal principle nonetheless know the sums required to determine the numerosity of collections of 6, 9 or 18 elements by breaking them up into small collections of 2 to 4 and then summing these. Therefore, it is unreasonable to assume that subset-knowers used this strategy. Thus, we believe that, short of postulating new representational systems for which no evidence has been provided, the only explanation left is that children compared the large numerosities with the ANS. Nevertheless, in Experiment 3, we directly address the concern of a lack of ratio effect by including a harder-to-discriminate ratio.

Another possible concern is that children controlled how long the collections were presented to them, so that, consequently, they could have used counting instead of parallel individuation and/or the ANS to compare numerosities. Since small collections are easier to count than large ones, this could explain why children were more accurate on small comparisons. However, the fact that children who did not know the cardinal principle – i.e., non-knowers and 1-knowers – showed this pattern of results makes this very unlikely. Nevertheless, in Experiment 3, we address this concern by asking children to compare the numerosities of collections that are presented too quickly to be counted.

Finally, it is possible that the better performance on comparisons of small numerosities was driven by better performance on the 1 vs. 3 comparisons only. By physical necessity, whenever one of the choices in a comparison is one object, the other choice is always the right answer. Thus, comparisons where one of the choices is a single object may be easier than all other comparisons. That is, one may predict that the effect of Number Range is specific to comparisons with a 1:3 ratio. Nevertheless, the lack of an interaction between Number Range and Ratio has ruled out this possibility.

There remain two alternatives that cannot be ruled out directly by the results of Experiment 1. First, it may be that children performed better on comparisons of small numerosities because the correct answer in this condition was always the same – i.e., 3 – whereas the correct answer in the large numerosity condition alternated between 9 and 18. Second, it could be that children were, in fact, relying on the ANS for all comparisons but that they performed more poorly on large comparisons because these comparisons make greater demands on non-numerical aspects of processing. For example, larger collections necessarily require children to divide their attention over more objects than small collections, and this could cause a decrement in performance. Although possible in theory, this “processing demands” alternative is unlikely to be the right explanation. On this alternative, the difficulty of comparisons should increase continuously as a function of the absolute size of numerosities. Contrary to this prediction, previous research provides strong evidence that when infants (Wood & Spelke, 2005) or adults (Barth, Kanwisher, & Spelke, 2003) use the ANS to compare numerosities, their level of performance for a given ratio of numerosities is the same regardless of the absolute size of the numerosities. Experiment 2 was explicitly designed to test both of these alternatives directly.

## Experiment 2

Like Experiment 1, Experiment 2 included one group of children who compared small numerosities only and another group who compared large numerosities only. Unlike Experiment 1, Experiment 2 only included size-number incongruent trials. This feature was changed to increase the number of trials that test whether children use distinct systems to compare collections on the basis of numerosity.

Experiment 2 had two main goals. First, we sought to replicate the effect of Number Range observed in Experiment 1. Second, we aimed to address the alternative explanations raised in Experiment 1. To address the processing demands alternative the large comparisons included a wider range of pairs of numerosities of the same ratio. To ensure that the number of trials was reasonable for young preschoolers, we only included comparisons with a ratio of 2:3, namely, 6 vs. 9, 10 vs. 15 and 12 vs. 18. Evidence that children’s accuracy decreases as numerosity increases would support this alternative. On the other hand, evidence that (1) children perform better when they compare small numerosities than when they compare large numerosities at the same ratio, and that (2) they perform equally well on all comparisons of large numerosities would provide strong evidence against the processing demands alternative.

We also tested whether the children in Experiment 1 who compared small numerosities were more accurate than those who compared large ones because the correct answer was always the same numerosity in the former condition but not in the latter. To test this, the small numerosity condition of Experiment 2 included comparisons of 1 vs. 2 and 3 vs. 4 in addition to comparisons of 2 vs. 3. Evidence that children were nonetheless more accurate on 2 vs. 3 than on the large comparisons would show that this alternative is false.

## Method

### Participants

A total of 86 2 ½ - to 4 ½-year-olds participated in this study. There were 45 children in the small numerosity condition, with an average age of 3 years 5 months (range: 2 years 6 months – 4 years 5 months; 24 males), and 41 children in the large numerosity condition, with an average age of 3 years 5 months (range: 2 years 6 months – 4 years 9 months; 21 males). A majority of the children were recruited in Southwestern Ontario, Canada ( $n = 63$ ; 24 in the small numerosity condition), and the remaining were recruited in the Greater Boston Area in the US ( $n = 23$ ; 21 in the small numerosity condition). Participants were predominantly monolingual speakers of English. An additional five children were excluded for always choosing the same side ( $n = 2$ ) and for failing to make a response ( $n = 3$ ).

### Design and Procedure

The stimuli and procedure were similar to Experiment 1. All children completed the numerical comparison task before the Give-*N* task, except one child who did not complete Give-*N*. The numerical comparison task had two between-subject conditions that differed on the range of numerosities tested: small numerosities ( $\leq 4$ ) and large ones ( $> 4$ ). As in Experiment 1, each individual collection was composed of red rectangles of the same size, and all collections were printed on letter-sized paper. Both ranges included pairs that differed by a 2:3 ratio: children in the small numerosity condition were asked to compare 2 vs. 3 (8 trials), and those in the large numerosity condition compared 6 vs. 9, 10 vs. 15, and 12 vs. 18 (4 trials each). To equate the total number of trials in the small and large conditions, the small numerosity condition also included comparisons of 1 vs. 2 (2 trials) and 3 vs. 4 (2 trials). Size and number were incongruent in all comparisons (as in Figures 1b and 1d). Pairs of collections of rectangles were presented in one of two item orders. The correct side (left or right) was pseudo-randomized such that no two consecutive trials were of the same pairs of numerosities. No feedback was given.

In the small numerosity condition, the cumulative surface area ranged from 6 to 8 cm<sup>2</sup> and cumulative perimeter ranged from 18 to 24 cm. In the large numerosity condition, cumulative surface area ranged from 18 to 36 cm<sup>2</sup>, and cumulative perimeter ranged from 54 cm to 108 cm.

## Results

### Give-*N*

The number of children and the mean age in each knower-level group are presented in Table 2.<sup>iv</sup>

Table 2

*Age and Number of Children in Each Knower-Level Group in Experiment 2*

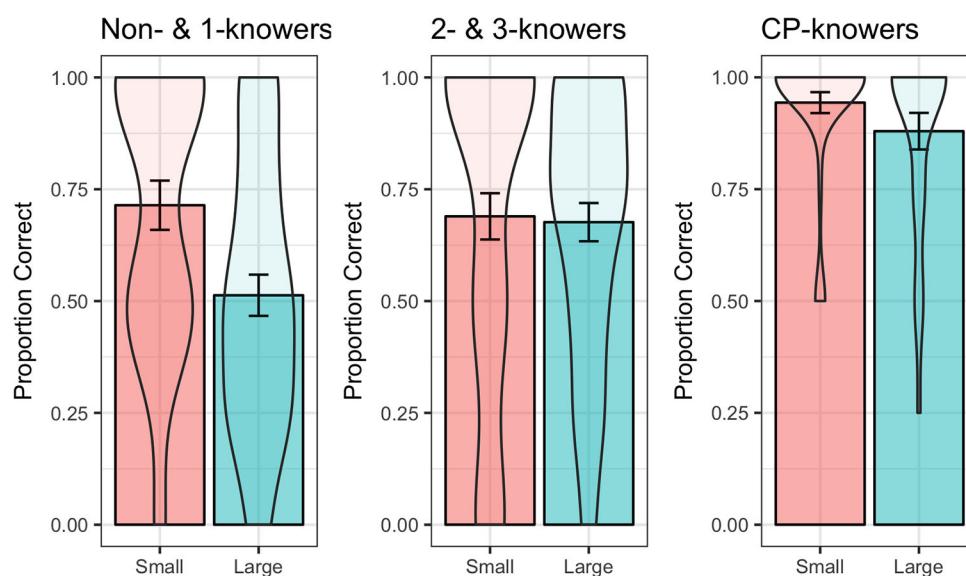
Knower-Level	Small			Large		
	<i>n</i>	Mean Age	Range	<i>n</i>	Mean Age	Range
Non-knowers	3	3;0	2;10 - 3;3	3	2;10	2;8 - 3;0
1-knowers	8	3;0	2;6 - 3;10	12	3;1	2;7 - 4;9
2-knowers	10	3;2	2;7 - 4;5	12	3;3	2;6 - 4;0
3-knowers	9	3;5	2;10 - 4;2	5	3;8	2;8 - 4;3
CP-knowers	14	3;10	3;3 - 4;5	9	4;1	3;4 - 4;7

Age was significantly correlated with children's knower-level, Pearson's  $r = .61$ ,  $p < .001$ .

### Numerical Comparisons

Preliminary analyses showed that children recruited in Canada and in the US performed similarly on the numerical comparison task,  $t(84) = -.28$ ,  $p = .78$ . Since almost all American children (21/23) were tested in the small numerosity condition, we also compared Americans ( $n = 21$ ) to Canadians ( $n = 24$ ) on comparisons of small numerosities only, and again, found no difference,  $t(43) = .30$ ,  $p = .77$ . Thus, we combined data from the two locations in subsequent analyses. Preliminary analyses also revealed that there were no order or gender effects, so we collapsed across these variables ( $ts < 1.4$ ,  $ps > .18$ ).

First, we asked if the effect of Number Range in Experiment 1 could be replicated. We constructed a linear regression using overall proportion correct on comparisons that differed by a 2:3 ratio as the dependent variable, and Number Range (small vs. large), Knower-Level (non- and 1-knowers, 2- and 3-knowers, and CP-knowers; CP-knowers as the reference category), and centered Age as independent variables. We found that adding the 3-way interaction or the 2-way interactions did not improve the fit of the regression model. We thus used a main-effects-only model in our final analysis. Results revealed a main effect of Number Range,  $\beta = -.13$ ,  $SE = .051$ ,  $p = .016$ . As in Experiment 1, children were better at comparing small numerosities ( $M = .81$ ,  $SD = .26$ ) than large numerosities ( $M = .66$ ,  $SD = .27$ ). There was also an effect of Age,  $\beta = .12$ ,  $SE = .052$ ,  $p = .019$ . Finally, we found a main effect of Knower-Level, with CP-knowers ( $n = 23$ ,  $M = .94$ ,  $SD = .14$ ) performing significantly better than 2- and 3-knowers,  $\beta = -.14$ ,  $SE = .069$ ,  $p = .045$  ( $n = 36$ ,  $M = .71$ ,  $SD = .28$ ), and non- and 1-knowers,  $\beta = -.20$ ,  $SE = .083$ ,  $p = .017$  ( $n = 26$ ,  $M = .60$ ,  $SD = .26$ ). Figure 3 displays children's performance on small and large numerosity comparisons.



**Figure 3.** Overall proportion correct on small and large numerosity comparisons on incongruent trials, separated by knower-levels in Experiment 2. Only 2:3 ratio comparisons were included. Violin plots are used to depict the distribution and probability density of the data. Error bars represent standard errors of the mean.

We also asked whether children succeeded on 2 vs. 3 because they somehow defaulted to choosing 3 or to avoiding 2 as the correct answer without doing the comparison. If children used the first strategy, they should

have always chosen the wrong collection on comparisons of 3 vs. 4. If they used the second, they should have always chosen the wrong collection on comparisons of 1 vs. 2. This was not so. Rather, children performed significantly above chance on 1 vs. 2 ( $M = .77$ ,  $SD = .38$ ),  $t(43) = 4.75$ ,  $p < .001$ , and on 3 vs. 4 ( $M = .75$ ,  $SD = .33$ ),  $t(43) = 5.0$ ,  $p < .001$ .

To test the processing demands alternative, we compared the large numerosity comparisons to each other. No significant differences were found (all  $p > .64$ ). We also compared 2 vs. 3 to each of the large numerosity comparisons individually. Multiple comparisons were corrected using the Holm-Bonferroni method. A significant difference was found in each case: 6 vs. 9,  $t(84) = 2.42$ ,  $p = .018$ ; 10 vs. 15,  $t(84) = 2.25$ ,  $p = .027$ ; 12 vs. 18,  $t(84) = 2.58$ ,  $p = .012$  (see Figure 4). These results strongly suggest that the Number Range effect was categorical – i.e., for all comparisons at a 2:3 ratio, children were more accurate on comparisons of small numerosities than on comparisons of large ones, but were equally accurate on all comparisons of large numerosities.

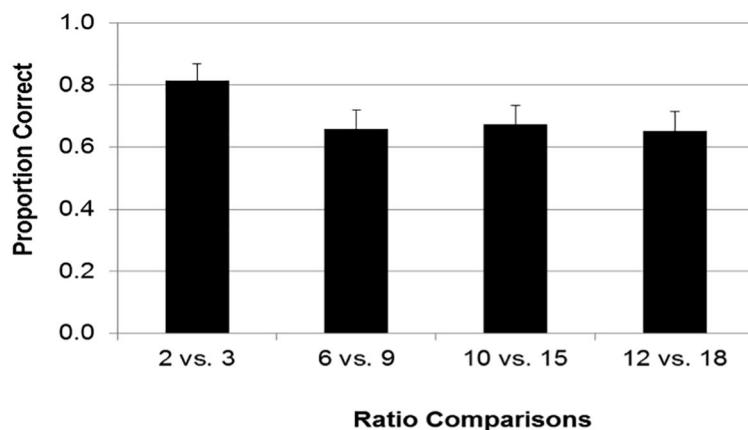


Figure 4. Proportion correct on small and each of the large numerosity comparisons in Experiment 2. Error bars represent standard error of the mean.

Finally, we asked whether children's ability to recruit parallel individuation to compare small numerosities is available even in children at the earliest stages of number word learning – i.e., non-knowers and 1-knowers. We found that they performed significantly above chance on small comparisons,  $t(10) = 2.64$ ,  $p = .013$  (1-tailed). We also examined whether non- and 1-knowers performed better on comparisons of small ( $M = .71$ ,  $SD = .27$ ) than large numerosities ( $M = .51$ ,  $SD = .22$ ), and found that they did,  $t(24) = 2.10$ ,  $p = .046$ ,  $d = .83$  [-0.02, 1.68]. However, the confidence interval is relatively large and includes 0, and thus, we should remain cautious when interpreting this finding. These data tentatively suggest that the development of the ability to use parallel individuation to compare numerosities likely does not depend on number word learning. These findings also suggest that children likely did not solve the numerical comparison task by counting.

## Discussion

Experiment 2 replicates several of the most important results of Experiment 1. First, it shows that children were more accurate on comparisons of small than large numerosities. We also found that non- and 1-knowers were significantly more accurate on small than on large comparisons. Since these children did not know the cardinal principle, it cannot be that this small-large difference was due to counting. In addition, counting is a poor explanation of the fact that children were more accurate on 2 vs. 3 than on the large comparisons, but were

equally accurate on all the latter comparisons. Indeed, if children had used counting to compare numerosities, one might have expected their performance to decrease gradually as numerosity increased.

Experiment 2 also rules out a number of alternative explanations of the fact that children were more accurate on comparisons of small numerosities. First, given that the numerosity of the correct answer varied in both the small and the large comparison conditions, the results of Experiment 2 show that greater accuracy on small comparisons in Experiment 1 was not due to the fact that the correct answer was always the same in the small numerosity condition but varied in the large numerosity condition. It also shows that children did not simply default to always choosing the collection of 3 or to avoiding the collection of 2 on the 2 vs. 3 comparison. Second, Experiment 2 provides evidence against what we called the processing alternative – i.e., they show that, contrary to the prediction of this alternative, children were more accurate on 2 vs. 3 than on comparisons of large numerosities of the same ratio, despite the fact that the large numerosities varied in size from 6 vs. 9 to 12 vs. 18.

## Experiment 3

The use of a between-subject design in the previous experiments leaves open the question of whether the effect of Number Range can be observed in the same children on the same task. To address this, we adopted a within-subject design, with the same group of children comparing exclusively small collections (2 vs. 3) and exclusively large collections (6 vs. 9). The small and large numerosity comparisons were presented in blocks in a counterbalanced order.

Experiment 3 made two further changes. First, although the fact that children who did not know the cardinal principle (e.g., non-knowers and subset-knowers) showed better performance on small collections than large collections suggests that this difference is not due to counting, we sought convergent evidence against this possibility by controlling the presentation time of the stimuli. That is, to prevent children from counting, the pairs of collections to be compared were presented for 2 seconds only (see also Abreu-Mendoza et al., 2013; Odic, Pietroski, et al., 2013 for similar presentation times). Previous studies suggest that 5- and 6-year-olds take, on average, 0.71 seconds (Trick, Enns, & Brodeur, 1996) to a little over 1 second (Chi & Klahr, 1975) to count an item.<sup>v</sup> Thus, to compare collections of 2 to collections of 3 by counting both of them, children would need at least 3.5 seconds, which exceeds the presentation time used in Experiment 3. Second, to ensure that our task engaged the ANS, we tested whether we could obtain a ratio effect in the large number range by including pairs of numerosities that differed by a 5:6 ratio, in addition to pairs that differed by a 2:3 ratio. Evidence that children perform more poorly on the 5:6 than on the 2:3 comparisons of large numerosities would confirm that our task taps the ANS.

As in Experiment 2, we only included size-number incongruent stimuli. We also included comparisons of 3 vs. 4 in the small numerosity block to prevent children from using the strategy of always choosing 3 without doing the comparison, and to match the small and large numerosity blocks in terms of number of types of comparisons (both blocks have two types).

## Method

### Participants

A total of 50 2 ½ - to 4 ½-year-olds participated in this study, with an average age of 3 years 5 months (range: 2 years 3 months – 4 years 8 months; 21 males, 29 females). An additional 18 children were tested and removed for providing no response ( $n = 4$ ), for failing to finish more than one block of trials ( $n = 5$ ), for always choosing the same side ( $n = 5$ ), for making a response before the stimuli appeared ( $n = 2$ ), and for parental interference ( $n = 1$ ). Twenty-five children completed the small comparison block first and 25 completed the large comparison block first. Children were recruited at a children's museum in Middletown, CT. No language information was collected from participants, but the population from which the sample was drawn from consists primarily of monolingual English speakers (78.6%; US Census Bureau).

### Design and Procedure

The procedure was similar to Experiment 2. All children completed the numerical comparison task before the Give-N task, except two children who did not complete Give-N. The numerical comparison task had two within-subject conditions that differed on the range of numbers tested: small numbers ( $\leq 4$ ) and large numbers ( $> 4$ ), presented in a counterbalanced order across children. As in Experiment 1, each individual collection was composed of rectangles of the same size. All children were first tested on four familiarization trials (1 vs. 3 and 10 vs. 30, two trials each).

The numerical comparison task always started with an experimenter introducing two stuffed animals - a blue bear sitting on the left side of a laptop computer and a red bear sitting on the right side. To ensure that children could identify the bears by their color, she asked children to point to the red bear and the blue bear. All children correctly identified the bears. Then, the experimenter said they were going to play a game with some blocks on the computer, and that their job was to indicate which bear had more blocks. The experimenter explained that the blue bear has blue blocks (while gesturing to the left of the computer) and that the red bear has red blocks (while gesturing to the right). Children were told that the blocks would go away very quickly and thus they would have to look at the screen very carefully.

On each trial, the experimenter asked, "Who has more blocks?" right before a white asterisk on a grey background appeared on the screen for 1500 ms. Then, an array of blue blocks and an array of red blocks appeared simultaneously for 2000 ms. Each array of blocks was presented on a 10 cm x 6 cm white background. The arrays were 4.3 cm apart from each other. The overall background of the screen was grey (see [Figures 5a](#) and [5b](#)). The experimenter entered the child's response through an external keyboard attached to the laptop as soon as the child responded. No feedback was given. All stimuli were presented on a 13" Macintosh laptop via PsychoPy.

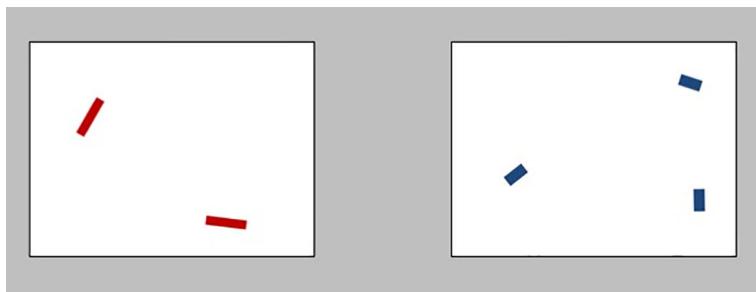


Figure 5a. A 2:3 ratio comparison in the small numerosity condition in Experiment 3 (size-number incongruent).

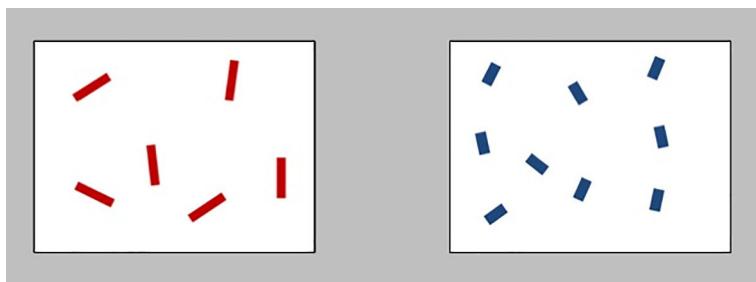


Figure 5b. A 2:3 ratio comparison in the large numerosity condition in Experiment 3 (size-number incongruent).

The study began with four familiarization trials, which included two collections of 1 vs. 3 and two collections of 10 vs. 30, presented in alternating order (1 vs. 3, 10 vs. 30, 1 vs. 3, 10 vs. 30). If a child failed to respond on the first trial, the experimenter repeated the instructions. If the child continued to provide no response, the experimenter terminated the study session and thanked the child for participating.

After the familiarization trials, the small and large comparisons were presented in blocks that were counterbalanced across children: half of the children completed the small comparisons block first and the other half completed the large comparisons block first. In the small comparisons block, children were asked to compare 2 vs. 3 (3 trials) and 3 vs. 4 (3 trials). In the large comparisons block, they were asked to compare 6 vs. 9 (3 trials) and 15 vs. 18 (3 trials). Stimuli for the 2 vs. 3, 3 vs. 4, and 6 vs. 9 comparisons were chosen from a subset of the stimuli used in Experiment 2. Each collection was digitized to ensure that the two collections in each pair of comparison had an equal number of pixels. Stimuli for the 15 vs. 18 comparisons were created in the same way as number-size incongruent stimuli in Experiment 1, and were then digitized as images.

Size and number were incongruent in all comparisons. Pairs of collections of rectangles were presented in one of two item orders. For one order, the first comparison was at a 2:3 ratio in both small and large comparisons blocks, and for the other order, the first comparison was at a 3:4 ratio in the small comparisons block and a 5:6 ratio in the large comparisons block. The correct side (left or right) was counterbalanced in each item order. The trials within each block were randomized such that no two consecutive trials were of the same ratio of comparison.

## Results

### Give-N

The number of children and the mean age in each knower-level group are presented in Table 3.<sup>vi</sup> Age was significantly correlated with children's knower-level, Pearson's  $r = .70$ ,  $p < .001$ .

Table 3

*Age and Number of Children in Each Knower-Level Group in Experiment 3.*

Knower-Level	n	Mean Age	Range
Non-knowers	6	2;9	2;4 - 3;3
1-knowers	9	2;8	2;3 - 3;0
2-knowers	8	3;7	2;8 - 4;2
3-knowers	5	3;9	3;6 - 4;0
CP-knowers	22	3;11	2;10 - 4;8

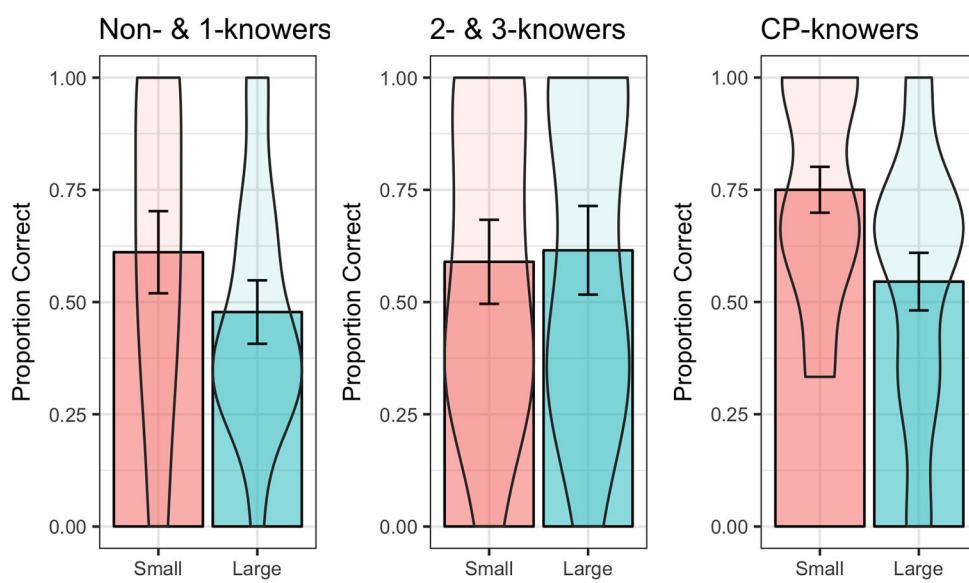
### Numerical Comparisons

Preliminary analyses revealed no effect of item order, block order, or gender, and thus, we collapsed across these variables in subsequent analyses (all  $ts < 1$ ,  $ps > .72$ ).

Children performed significantly above chance on the practice trials,  $t(49) = 5.84$ ,  $p < .001$ ,  $M = .71$ ,  $SD = .26$ , suggesting that, despite the short presentation time, they had no difficulty understanding the instructions and perceiving the arrays. Children performed similarly on the small (1 vs. 3;  $M = .71$ ,  $SD = .37$ ) and the large ( $M = .71$ ,  $SD = .30$ ) comparisons at a 1:3 ratio.

To examine the effect of number range, we constructed a logistic mixed-effects model predicting children's correct responses on 2:3 ratio comparisons (2 vs. 3 and 6 vs. 9) with Number Range (small vs. large), centered Age in months, and Knower-Level (non- and 1-knowers, 2- and 3-knowers vs. CP-knowers) as fixed factors, with by-subject random slopes for Number Range. We also included 2-way interaction terms involving Number Range (Number Range x Age, Number Range x Knower-Level). This model did not increase the fit over a model with main effects,  $\chi^2(3) = 3.21$ ,  $p = .36$ , and we thus used a main-effects-only model in our final analysis. Replicating the previous two experiments, we found a main effect of Number Range,  $\beta = -.56$ ,  $SE = .27$ ,  $z = -2.11$ ,  $p = .035$ . Children were better at comparing small numerosities ( $M = .67$ ,  $SD = .31$ ) than large numerosities ( $M = .54$ ,  $SD = .31$ ). We also found an effect of Age,  $\beta = .085$ ,  $SE = .026$ ,  $z = 3.33$ ,  $p < .001$ . No other effects were found. Figure 6 displays children's performance on small and large comparisons by knower-levels.

Next, we asked if there was a ratio effect in the large number range. To test this, we selected children who answered at least two out of three 6 vs. 9 comparisons correctly and asked if they performed worse on the 15 vs. 18 comparisons ( $n = 26$ ,  $M_{age} = 44.1$  months). We found that these children did perform better on 6 vs. 9 ( $M = .79$ ,  $SD = .17$ ) than on 15 vs. 18 comparisons ( $M = .49$ ,  $SD = .29$ ),  $t(25) = -4.63$ ,  $p < .001$ ,  $d_{av} = 1.30$  [.44, 1.36]<sup>vii</sup>, indicating that the comparison with a 2:3 ratio comparison was easier than the one with a 5:6 ratio. Nevertheless, children who could compare 6 vs. 9 were no better at comparing 10 vs. 30 ( $M = .75$ ,  $SD = .32$ ),  $t(25) = -.63$ ,  $p = .54$ ,  $d_{av} = 0.16$  [-.51, .27].



*Figure 6.* Overall proportion correct on small and large numerosity comparisons on incongruent trials, separated by knower-levels in Experiment 3. Only 2:3 ratio comparisons were included. Violin plots are used to depict the distribution and probability density of the data. Error bars represent standard errors of the mean.

We also asked whether children succeeded on 2 vs. 3 because they somehow defaulted to choosing 3 without doing the comparison. Against this explanation, children performed significantly above chance on 3 vs. 4 ( $M = .63$ ,  $SD = .29$ ),  $t(49) = 3.08$ ,  $p = .003$ .

Unlike Experiments 1 and 2, Experiment 3 did not show a main effect of knower-level – CP-knowers ( $M = .65$ ,  $SD = .29$ ) were not significantly more accurate than 2- and 3-knowers ( $M = .54$ ,  $SD = .34$ ) or non- and 1-knowers ( $M = .60$ ,  $SD = .32$ ). However, a marginally larger proportion of CP-knowers answered at least 4 out of 6 trials correctly (16/22; 72.7%) compared to 2- and 3-knowers (8/13; 61.5%) and non- and 1-knowers (5/15; 33.3%;  $\chi^2(2) = 5.77$ ,  $p = .056$ ), suggesting that, although the difference may have been small, CP-knowers were better at comparing numerosities non-verbally than subset-knowers and non-knowers.

Finally, we asked whether children who were at the earliest stages of number word learning used parallel individuation to compare small numerosities. Contrary to what we found in Experiments 1 and 2, we found that non-knowers and 1-knowers did not perform significantly above chance on 2 vs. 3 ( $M = .61$ ,  $SD = .35$ ),  $t(14) = 1.21$ ,  $p = .12$  (1-tailed). We also found that they did not perform significantly better on 2 vs. 3 than on 6 vs. 9 comparisons ( $M = .48$ ,  $SD = .27$ ;  $t(14) = 1.13$ ,  $p = .28$ ,  $d_{av} = .42$  [-.23, .81]). It may be that Experiment 3 did not replicate this aspect of the results of Experiments 1 and 2 because the use of short presentation times reduced the size of the difference between non-knowers and 1-knowers' accuracy on comparisons of small numerosities and their accuracy on comparisons of large ones. That is, while the greater difficulty of the comparison task in Experiment 3 could reduce non-knowers' and 1-knowers' performance for small comparisons, it could not reduce it for large comparisons because their performance on these comparisons was already at floor (0.5) in Experiments 1 and 2. Consequently, the greater difficulty of the task in Experiment 3 might have reduced the size of the difference between accuracy on small and large comparisons for this group of children. Indeed, the effect size observed here (.42; mean small-large difference = .13) is about half as small as those observed in Experiments 1 (.77; mean difference = .22) and 2 (.83; mean difference = .20).

### A Mini Meta-Analysis of Children at the Earliest Stages of Number Word Learning

Although non-knowers and 1-knowers performed better on small comparisons than large ones in all three experiments, the mean difference was significant in only two of the three experiments. We thus sought to gather cumulative evidence on whether non- and 1-knowers can recruit parallel individuation to compare small collections by conducting a mini meta-analysis (see [Goh, Hall, & Rosenthal, 2016](#) for discussion). We first used a fixed effects approach in which the effect sizes were weighted by sample size. We then converted Cohen's  $d$  into Pearson's  $r$  ( $r = .36, .38$ , and  $.21$ , respectively for Experiments 1, 2 and 3), and transformed it to Fisher's  $z$  scores. Overall, the difference between small and large comparisons for non- and 1-knowers was significant,  $d = .77$ ,  $Z = 2.85$ ,  $p = .0044$ . A random effects approach averaging across the effect sizes also revealed a similar finding. The effect size was significantly different from 0,  $d = .67$ ,  $t(2) = 5.27$ ,  $p = .034$ . These analyses suggest that it is likely that children at the earliest stages of number word learning can recruit parallel individuation to support numerical comparisons.

## Discussion

Experiment 3 made two important methodological changes in an attempt to replicate and expand previous findings. We adopted a within-subject design such that the same children compared pairs of small and pairs of large numerosities, and we prevented counting by limiting presentation time to 2 seconds. Despite these changes, we replicated the finding that 2 vs. 3 comparisons were easier than comparisons of large numerosities with an equal ratio. This provides strong evidence that preschoolers can use parallel individuation to compare small numerosities.

The lack of a number range effect at the 1:3 ratio may raise doubts about whether the small-large difference reflects the use of distinct systems for each number range. However, to assess whether children can recruit parallel individuation (instead of the approximate number system) to compare small numerosities, one need not show that small numerosity comparisons are different from large numerosity comparisons *at all ratios*. The finding that they are different at some of the ratios is sufficient to support the claim that parallel individuation can support numerical comparisons. Consistent with our lack of accuracy difference at the 1:3 ratio, [Choo and Franconeri \(2014\)](#) have demonstrated that adults' RTs for comparing 1 vs. 3 and 10 vs. 30 were similar, but they were nevertheless different for comparing 2 vs. 3 and 20 vs. 30. They take their results to suggest that adults do not use the same systems when they compare small numerosities and when they compare large ones. Thus, we too take our results to be consistent with the view that children do not recruit the same systems when they compare small numerosities and when they compare large ones.

Finally, we found that CP-knowers were better at comparing numerosities non-verbally than subset-knowers and non-knowers. However, the difference between CP-knowers and subset- and non-knowers observed in Experiment 3 was smaller than in Experiments 1 and 2.

Unlike what was found in Experiments 1 and 2, we found that children who were at the earliest stages of number word learning did not show a significant number range effect. However, a mini meta-analysis summarizing across all three experiments suggests that the difference between small and large numerosity comparisons is likely a reliable effect for non- and 1-knowers.

## General Discussion

In three different experiments, we find that preschoolers between the ages of two-and-a-half and four-and-a-half – were modestly, but reliably more accurate in comparing collections of 2 and 3 than pairs of numerosities larger than 5 with the same ratio. This difference was found regardless of whether performance on comparisons of small numerosities was compared to performance on large numerosities across different children or within the same ones, and regardless of whether the presentation time was child-controlled or limited to prevent counting. This suggests that it is a highly robust result.

On our view, the best explanation of this difference is that, in this task, children use distinct systems to compare small and large numerosities: parallel individuation for 1 to 3 and the ANS for more than 5. Various aspects of our results support this conclusion. We know that children based their choices on the numerosities of the collections because the evidence for this explanation was found on comparisons where size and numerosity were incongruent – i.e., where children could not base their decision on the total area or the total perimeter of the collections because these were equated, and where they could not base it on the individual object sizes because these were larger in the collection with the smaller numerosity. Our results cannot be explained by counting because children were more accurate on small than on large comparisons even if (1) they did not understand the cardinal principle and thus could not count to compare numerosities and (2) they could not count any of the collections because they were presented too quickly to allow them to do so. Moreover, it is unlikely that children were more accurate on comparisons of small than large numerosities because they could name the former but not the latter. Indeed, in two out of three experiments, we found the same pattern of results in children who could not name the small numerosities – namely non-knowers and 1-knowers. Our results cannot be explained by any hypothesis that predicts that accuracy decreases gradually as numerosity increases – i.e., while children were more accurate on comparisons of small numerosities (2 vs. 3) than on any of the comparisons of large numerosities, they were equally accurate on comparisons of a relatively wide range of large numerosities (6 vs. 9, 10 vs. 15, and 12 vs. 18). Finally, Experiment 3 provided direct evidence that our task did engage the ANS when numerosities were large – i.e., children performed more poorly on comparisons of 15 vs. 18 than on comparisons of 6 vs. 9, suggesting that their accuracy on these comparisons was controlled by the ratio of the numerosities and not by the absolute difference between them. This leaves the view that children used parallel individuation for small comparisons and the ANS for large ones as the only plausible explanation of our results.

To be clear, we are not claiming that the ANS is never used to represent small numerosities. Indeed, some studies have shown that, in some contexts, infants (Cordes & Brannon, 2009; Starr, Libertus, & Brannon, 2013b) and adults (Burr, Turi, & Anobile, 2010; Cordes et al., 2001; Hyde & Wood, 2011) use the ANS to represent small numerosities. We do not believe that the present results conflict with these studies. Rather, our results show that parallel individuation *can* be used to compare numerosities of two distinct collections but they do not mean that it is the *only* system that can do so. While our studies were not designed to say why children used parallel individuation instead of the ANS to compare the numerosities of the particular collections they saw, we presume that the explanation has to do with factors that were identified in studies that were designed for that purpose – e.g., whether participants have access to their full attentional resources, and whether the spacing between individual objects is within the spatial resolution of attention (see Burr, Turi, & Anobile, 2010; Hyde, 2011; Hyde & Wood, 2011; Starr et al., 2013b).

## Can Children Use Parallel Individuation to Compare Small Numerosities at the Earliest Stages of Number Word Learning?

Across all three experiments, we found that children who had not learned the meaning of any number word beyond “one” (i.e., non- and 1-knowers) demonstrated better performance on small numerosity comparisons than on large numerosity comparisons. Although the difference between small and large numerosity comparisons was significant in only two of the three experiments, a mini meta-analysis using effect sizes suggests that the number range effect in this group of children is reliable. Thus, although the difference between performance on small and large comparisons may be small for this group, we take it to suggest that it is likely that the capacity of parallel individuation to support numerical comparisons is available prior to the acquisition of number word meanings.

Interestingly, studies of comparison of small collections in infants suggest that it might not be available in the form observed here from the beginning of life. On the one hand, Feigenson (2005) finds that, when each element in collections of fewer than 4 elements has a unique color, pattern and texture (Feigenson, 2005), young infants can discriminate between collections of fewer than 4 on the basis of their numerosity. However, it is not possible to infer with certainty that the infants in Feigenson’s (2005) study used parallel individuation to compare the collections because she did not compare infants’ performance on small collections to their performance on large ones. Still, her data are consistent with this possibility. On the other hand, other studies of numerical discrimination in infants suggest that they do not distinguish collections of fewer than four elements on the basis of numerosity when all the elements have the same properties (e.g., Feigenson, Carey, & Spelke, 2002; Xu, 2003; Xu et al., 2005). Moreover, Feigenson, Carey, and Hauser (2002) report that when 10 and 12-month-olds are offered a choice between two small numerosities of crackers, they base their decision on the total amount of cookie instead of numerosity. Therefore, it may be that young infants can use parallel individuation to compare collections on the basis of numerosity but that they do so in a narrower range of contexts than preschoolers. Whether the differences between infants and preschoolers reflect genuine cognitive changes or merely differences in the paradigms used with the two populations ought to be investigated in future studies.

## Relation Between Knowledge of the Cardinal Principle and Performance on Non-Verbal Numerical Comparisons

In two of the three experiments, we found that CP-knowers were significantly better at numerical comparisons than subset-knowers. In the last experiment, we did not find a significant effect of acquiring the cardinal principle on average accuracy, but we found that CP-knowers were more likely to answer at least 4 out of 6 trials correctly than subset-knowers and non-knowers. Two other studies of numerical comparisons where the relative size of the elements in the collections conflicts with their relative numerosity (what we called “size-number incongruent” comparisons) show similar results (Negen & Sarnecka, 2015; Wagner & Johnson, 2011; see also Abreu-Mendoza et al., 2013). Negen and Sarnecka report that poor performance in this age is specifically due to subset-knowers – they perform at chance, whereas CP-knowers perform above chance. Moreover, although Wagner and Johnson administered a modified version of the Give-*N* task and did not analyze subset-knowers and CP-knowers separately, inspection of their data suggests a similar pattern of results, with a majority of children who could generate sets larger than 4 on the Give-*N* task performed above 50% and those who could only

generate small sets performed at around 50%. Thus, studies conducted by three separate laboratories have now shown this effect.

Why is acquiring the cardinal principle related to an improvement in children's performance on non-verbal numerical comparisons? On one view, the acquisition of the cardinal principle causes the representations of numerosity created by the ANS to become more precise (Shusterman, Slusser, Halberda, & Odic, 2016). However, Negen and Sarnecka (2015) provide evidence against this hypothesis and show that it is possible for subset-knowers to perform like CP-knowers. They presented children aged between two-and-a-half and six years with comparisons of pairs of large numerosities ( $> 20$ ) where the sum of the areas of the individual elements comprising the numerically larger collection was smaller than that of the numerically smaller collection. They replicated previous findings, showing that subset-knowers performed at chance on numerical comparisons and that CP-knowers performed above chance. Importantly, they also found that after subset-knowers were trained to choose the numerically larger collection (instead of the collection with larger individual objects and larger total area) in comparisons at a 1:3 ratio, they were able to generalize this training to comparisons at other ratios, and performed at the same level as CP-knowers.

In light of these results, Negen and Sarnecka argue that it is not the case that the acquisition of the cardinal principle provokes changes in the precision of the ANS. Rather, they suggest that (1) subset-knowers still do not clearly differentiate the continuous meaning of "more" from its discrete meaning (as in "more chocolate" vs. "more candy bars"), and that (2) CP-knowers are more likely than subset-knowers to spontaneously pay more attention to numerosity than to area. Our data provide some evidence against the first part of Negen and Sarnecka's hypothesis. We found that in Experiments 1 and 2, children who had not acquired the cardinal principle (i.e., non-, 1-, 2-, and 3-knowers) performed above chance on comparisons of small numerosities ( $p's < .001$ ), even when size and numerosity were incongruent. This suggests that they can distinguish between the continuous and the discrete meaning of the word "more". Thus, we believe that the best explanation is one that goes along the lines of the second part of Negen and Sarnecka's hypothesis, namely that, when asked to compare the numerosities of two collections, the relative physical dimensions of the collections such as the size of their elements and the sum of their areas interfere with relative numerosity, and subset-knowers cannot resolve this conflict when the ANS is the only system that can represent the numerosity of the collections (i.e., when both collections contain more than 4 elements) whereas CP-knowers can (see Abreu-Mendoza et al., 2013 for a similar proposal).<sup>viii</sup>

## How Does Parallel Individuation Support Numerical Comparison of Distinct Collections?

Unlike the ANS, the symbols created with parallel individuation are not symbols for cardinalities; they are symbols for objects. For example, when infants use parallel individuation to represent two toy cars hidden in a box, they do not represent the contents of the box as *two cars* but rather as *a car and another car* or *an object and another object* (Feigenson & Carey, 2003, 2005). Given the nature of the representation, how does parallel individuation support the computation of numerical comparisons? Although we can only speculate, we suggest that a natural possibility is that it does so via computations of one-to-one correspondence. We know from work by Feigenson and colleagues (Feigenson, Carey, & Hauser, 2002; Feigenson & Halberda, 2008; Rosenberg & Feigenson, 2013) that infants can re-organize representations of objects created with parallel individuation into "chunks." This allows them to keep track of two separate collections, with a maximum capacity of two objects in

one collection and three in the other. Thus, it may be that, when young preschoolers use parallel individuation to solve numerical comparisons, they create a representation of the elements of each collection to be compared (e.g., “a red block” vs. “a blue block and another blue block”) and determine which collection is more numerous by comparing the two collections on the basis of one-to-one correspondence (e.g., the red block matches with a blue block; one blue block is left unmatched; therefore, there are more blue blocks). Future studies should investigate whether this is indeed the mechanism whereby parallel individuation supports numerical comparisons.

## Conclusions

The present study is the first to provide evidence that the ability to use parallel individuation to compare distinct collections on the basis of numerosity is available at least by age 2 and that the development of this ability likely does not depend on number word learning. Following Carey and colleagues (Carey, 2009; Feigenson & Carey, 2005; Le Corre & Carey, 2007), we suggest that the system compares numerosities by comparing collections on the basis of one-to-one correspondence. Thus, there may be two developmental roots of our knowledge of numerical relations: the one-to-one correspondence operations defined over representations of objects created with parallel individuation, and the comparisons of the size of the mental magnitudes created with the ANS.

## Notes

- i) We fit all mixed effects models using the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015). We used mixed effects models in Experiments 1 and 3 because the design of these studies violated the independence assumption of regression models, and thus, unlike in Experiment 2, these data could not be analyzed with linear regression.
- ii) As a stronger test of whether the capacity to use parallel individuation to compare numerosities depends on number word learning, we also analyzed non-knowers on their own – i.e., without 1-knowers. We found that the non-knowers performed significantly above chance on these comparisons,  $M = .66$ ,  $SD = .23$ ,  $t(7) = 1.93$ ,  $p = .047$  (1-tailed). They also performed marginally better on comparisons of small numerosities (1 vs. 3 and 2 vs. 3) than of large numerosities (6 vs. 18 and 6 vs. 9;  $M = .43$ ,  $SD = .24$ ;  $t(13) = 1.89$ ,  $p = .082$  (2-tailed),  $d = .98$ , 95% CI [-0.21, 2.16]).
- iii) All effect sizes (Cohen’s  $d$ ) and their 95% confidence interval for independent samples t-tests were computed using the compute.es package in R (AC Del Re, 2013). Effect sizes for paired-samples t-tests (Cohen’s  $d_{av}$ ) were computed using Lakens (2013)’s Supplementary Materials; their confidence interval was computed using an approximation strategy outlined in Howell (2009).
- iv) Across both small and large numerosity conditions, we found 1 4-knower and he was classified as a 3-knower in the analyses. One child in the small numerosity comparison did not complete the Give-N task but was included in the numerical comparison task analysis whenever knower-level was not a factor.
- v) Trick and colleagues (1996) reported that 6-year-old children took 4503 ms to count 6 items, 4887 ms to count 7 items, 5388 to count 8 items, and 6190 to count 9 items. Thus, the time 6-year-olds took to count one item ranges from approximately .67 ms to .75 ms.
- vi) We found one 4-knower. She was grouped with the 3-knowers in the analyses.
- vii) We reported Cohen’s  $d_{av}$  because it is comparable to Cohen’s  $d$  for within-subject designs (see Lakens, 2013 for discussion).
- viii) We conducted post-hoc analyses to further explore this proposal. Despite a lack of an interaction between Size-Number Congruence and Knower-level, using data from Experiment 1, we found that for large comparisons, subset-knowers were at chance on size-number incongruent trials ( $M = .55$ ,  $SD = .32$ ;  $t(32) < 1$ ,  $p = .41$ ) but above chance on size-number congruent trials ( $M = .77$ ,  $SD = .27$ ;  $t(32) = 5.6$ ,  $p < .001$ ). CP-knowers, on the other hand, were above chance on both types of trials ( $M = .86$  for incongruent and  $M = .97$  for congruent;  $ts > 3.25$ ,  $ps < .012$ ). Multiple comparisons were corrected using the Holm-Bonferroni method. Given that this analysis is exploratory, future studies should confirm this pattern of results.

## Funding

This research was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant 371718-2009 and a PRODEP grant to MLC and an NSERC CGS-D scholarship to PC.

## Competing Interests

The authors have declared that no competing interests exist.

## Acknowledgments

We thank all the children and families for participating, and preschools and teachers for their help with this project. Without them, this research would not be possible. We also want to thank Miranda Sollychin and Alexandra Butti for their assistance with data collection. We are grateful to Susan Carey and Anna Shusterman for allowing the first author to use their lab resources to recruit participants for Experiment 2 and Experiment 3, respectively. We thank two anonymous reviewers for providing helpful feedback on this manuscript.

## Data Availability

We cannot share our data on a public platform because we did not ask participants for their permission to do so on our consent forms. Nevertheless, to obtain the raw data, interested readers can contact Pierina Cheung at [cheung.pierina@gmail.com](mailto:cheung.pierina@gmail.com). For study materials and analysis code see the [Supplementary Materials](#).

## Supplementary Materials

All study materials and analysis code can be found on the Open Science Framework project page (see [Index of Supplementary Materials](#)).

### Index of Supplementary Materials

Cheung, P., & Le Corre, M. (2017). *Parallel individuation in preschoolers* [Supplementary materials]. <https://osf.io/dg4qr/>

## References

- Abreu-Mendoza, R. A., Soto-Alba, E. E., & Arias-Trejo, N. (2013). Area vs. density: Influence of visual variables and cardinality knowledge in early number comparison. *Frontiers in Psychology*, 4, Article 805.  
doi:[10.3389/fpsyg.2013.00805](https://doi.org/10.3389/fpsyg.2013.00805)
- AC Del Re. (2013). compute.es: Compute Effect Sizes (R package version 0.2-2) [Computer software]. Retrieved from <http://cran.r-project.org/web/packages/compute.es>
- Barner, D., & Snedeker, J. (2005). Quantity judgments and individuation: Evidence that mass nouns count. *Cognition*, 97, 41-66. doi:[10.1016/j.cognition.2004.06.009](https://doi.org/10.1016/j.cognition.2004.06.009)
- Barner, D., & Snedeker, J. (2006). Children's early understanding of mass-count syntax: Individuation, lexical content, and the number asymmetry hypothesis. *Language Learning and Development*, 2, 163-194.  
doi:[10.1207/s15473341lld0203\\_2](https://doi.org/10.1207/s15473341lld0203_2)
- Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, 86, 201-221. doi:[10.1016/S0010-0277\(02\)00178-6](https://doi.org/10.1016/S0010-0277(02)00178-6)

- Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Nonsymbolic arithmetic in adults and young children. *Cognition*, 98, 199-222. doi:[10.1016/j.cognition.2004.09.011](https://doi.org/10.1016/j.cognition.2004.09.011)
- Barth, H., La Mont, K., Lipton, J., & Spelke, E. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 14116-14121. doi:[10.1073/pnas.0505512102](https://doi.org/10.1073/pnas.0505512102)
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:[10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)
- Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition*, 83, 223-240. doi:[10.1016/S0010-0277\(02\)00005-7](https://doi.org/10.1016/S0010-0277(02)00005-7)
- Brannon, E. M., & Van de Walle, G. A. (2001). The development of ordinal numerical competence in young children. *Cognitive Psychology*, 43, 53-81. doi:[10.1006/cogp.2001.0756](https://doi.org/10.1006/cogp.2001.0756)
- Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision*, 10, Article 20. doi:[10.1167/10.6.20](https://doi.org/10.1167/10.6.20)
- Cantlon, J. F., Safford, K. E., & Brannon, E. (2010). Spontaneous analog number representations in three-year-old children. *Developmental Science*, 13, 289-297. doi:[10.1111/j.1467-7687.2009.00887.x](https://doi.org/10.1111/j.1467-7687.2009.00887.x)
- Cantrell, L., Kuwubara, M., & Smith, L. (2015). Set size and culture influence children's attention to number. *Journal of Experimental Child Psychology*, 131, 19-37. doi:[10.1016/j.jecp.2014.10.010](https://doi.org/10.1016/j.jecp.2014.10.010)
- Carey, S. (2009). *The origin of concepts*. New York, NY, USA: Oxford University Press.
- Chi, M. T. H., & Klahr, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology*, 19, 434-439. doi:[10.1016/0022-0965\(75\)90072-7](https://doi.org/10.1016/0022-0965(75)90072-7)
- Choo, H., & Franconeri, S. L. (2014). Enumeration of small collections violates Weber's Law. *Psychonomic Bulletin & Review*, 21, 93-99. doi:[10.3758/s13423-013-0474-4](https://doi.org/10.3758/s13423-013-0474-4)
- Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, 10, 408-411. doi:[10.1111/1467-9280.00177](https://doi.org/10.1111/1467-9280.00177)
- Clearfield, M. W., & Mix, K. S. (2001). Amount versus number: Infants' use of area and contour length to discriminate small sets. *Journal of Cognition and Development*, 2, 243-260. doi:[10.1207/S15327647JCD0203\\_1](https://doi.org/10.1207/S15327647JCD0203_1)
- Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review*, 8, 698-707. doi:[10.3758/BF03196206](https://doi.org/10.3758/BF03196206)
- Cordes, S., & Brannon, E. M. (2009). Crossing the divide: Infants discriminate small from large numerosities. *Developmental Psychology*, 45, 1583-1594. doi:[10.1037/a0015666](https://doi.org/10.1037/a0015666)
- Coubart, A., Izard, V., Spelke, E. S., Marie, J., & Streri, A. (2014). Dissociation between small and large numerosities in newborn infants. *Developmental Science*, 17, 11-22. doi:[10.1111/desc.12108](https://doi.org/10.1111/desc.12108)
- Dehaene, S. (1997). *The number sense*. New York, NY, USA: Oxford University Press.

- De Smedt, B., Noël, M., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, 2, 48-55. doi:[10.1016/j.tine.2013.06.001](https://doi.org/10.1016/j.tine.2013.06.001)
- Feigenson, L. (2005). A double dissociation in infants' representation of object arrays. *Cognition*, 95, B37-B48. doi:[10.1016/j.cognition.2004.07.006](https://doi.org/10.1016/j.cognition.2004.07.006)
- Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science*, 6, 568-584. doi:[10.1111/1467-7687.00313](https://doi.org/10.1111/1467-7687.00313)
- Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition*, 97, 295-313. doi:[10.1016/j.cognition.2004.09.010](https://doi.org/10.1016/j.cognition.2004.09.010)
- Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object-files versus analog magnitudes. *Psychological Science*, 13, 150-156. doi:[10.1111/1467-9280.00427](https://doi.org/10.1111/1467-9280.00427)
- Feigenson, L., Carey, S., & Spelke, E. S. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology*, 44, 33-66. doi:[10.1006/cogp.2001.0760](https://doi.org/10.1006/cogp.2001.0760)
- Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8, 307-314. doi:[10.1016/j.tics.2004.05.002](https://doi.org/10.1016/j.tics.2004.05.002)
- Feigenson, L., & Halberda, J. (2008). Conceptual knowledge increases infants' memory capacity. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 9926-9930. doi:[10.1073/pnas.0709884105](https://doi.org/10.1073/pnas.0709884105)
- Frank, M. C., & Barner, D. (2012). Representing exact number visually using mental abacus. *Journal of Experimental Psychology: General*, 141, 134-149. doi:[10.1037/a0024427](https://doi.org/10.1037/a0024427)
- Gallistel, C. R. (1989). Animal cognition: The representation of space, time and number. *Annual Review of Psychology*, 40, 155-189. doi:[10.1146/annurev.ps.40.020189.001103](https://doi.org/10.1146/annurev.ps.40.020189.001103)
- Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, 4, 59-65. doi:[10.1016/S1364-6613\(99\)01424-2](https://doi.org/10.1016/S1364-6613(99)01424-2)
- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., . . . Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS One*, 8(6), Article e67374. doi:[10.1371/journal.pone.0067374](https://doi.org/10.1371/journal.pone.0067374)
- Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass*, 10, 535-549. doi:[10.1111/spc3.12267](https://doi.org/10.1111/spc3.12267)
- Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society B: Biological Sciences*, 267, 829-833. doi:[10.1098/rspb.2000.1078](https://doi.org/10.1098/rspb.2000.1078)
- Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "Number Sense": The approximate number system in 3-, 4-, 5-, 6-year-olds and adults. *Developmental Psychology*, 44, 1457-1465. doi:[10.1037/a0012682](https://doi.org/10.1037/a0012682)
- Halberda, J., Mazzocco, M., & Feigenson, L. (2008). Individual differences in nonverbal number acuity correlate with maths achievement. *Nature*, 455, 665-669. doi:[10.1038/nature07246](https://doi.org/10.1038/nature07246)

- Howell, D. C. (2009). Confidence intervals on effect size [Supplemental material to Chapter 7]. In *Statistical methods for psychology* (7th ed.). Belmont, CA, USA: Cengage Wadsworth. Retrieved from <https://www.uvm.edu/~dhowell/methods7/Supplements/Confidence%20Intervals%20on%20Effect%20Size.pdf>
- Hyde, D. C. (2011). Hypothesis and theory article: Two systems of non-symbolic numerical cognition. *Frontiers in Human Neuroscience*, 5, Article 150. doi:[10.3389/fnhum.2011.00150](https://doi.org/10.3389/fnhum.2011.00150)
- Hyde, D. C., & Spelke, E. S. (2011). Neural signatures of number processing in human infants: Evidence for two core systems underlying numerical cognition. *Developmental Science*, 14, 360-371. doi:[10.1111/j.1467-7687.2010.00987.x](https://doi.org/10.1111/j.1467-7687.2010.00987.x)
- Hyde, D. C., & Wood, J. N. (2011). Spatial attention determines the nature of non-verbal numerical cognition. *Journal of Cognitive Neuroscience*, 23, 2336-2351. doi:[10.1162/jocn.2010.21581](https://doi.org/10.1162/jocn.2010.21581)
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, 4, Article 863. doi:[10.3389/fpsyg.2013.00863](https://doi.org/10.3389/fpsyg.2013.00863)
- 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. doi:[10.1016/j.cognition.2006.10.005](https://doi.org/10.1016/j.cognition.2006.10.005)
- Le Corre, M., & Carey, S. (2008). Why the verbal counting principles are constructed out of representations of small sets of individuals: A reply to Gallistel. *Cognition*, 107, 650-662. doi:[10.1016/j.cognition.2007.09.008](https://doi.org/10.1016/j.cognition.2007.09.008)
- Le Corre, M., Van de Walle, G., Brannon, E., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology*, 52, 130-169. doi:[10.1016/j.cogpsych.2005.07.002](https://doi.org/10.1016/j.cogpsych.2005.07.002)
- Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, 14, 1292-1300. doi:[10.1111/j.1467-7687.2011.01080.x](https://doi.org/10.1111/j.1467-7687.2011.01080.x)
- Lipton, J. S., & Spelke, E. S. (2004). Origins of number sense: Large number discrimination in human infants. *Psychological Science*, 14, 396-401. doi:[10.1111/1467-9280.01453](https://doi.org/10.1111/1467-9280.01453)
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279-281. doi:[10.1038/36846](https://doi.org/10.1038/36846)
- McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, 15, 776-781. doi:[10.1111/j.0956-7976.2004.00755.x](https://doi.org/10.1111/j.0956-7976.2004.00755.x)
- McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science*, 18, 740-745. doi:[10.1111/j.1467-9280.2007.01969.x](https://doi.org/10.1111/j.1467-9280.2007.01969.x)
- Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, 9, 320-334. doi:[10.1037/0097-7403.9.3.320](https://doi.org/10.1037/0097-7403.9.3.320)
- Moher, M., Tuerk, A. S., & Feigenson, L. (2012). Seven-month-old infants chunk items in working memory. *Journal of Experimental Child Psychology*, 112, 361-377. doi:[10.1016/j.jecp.2012.03.007](https://doi.org/10.1016/j.jecp.2012.03.007)
- Moyer, R. S., & Landauer, T. K. (1967). Time required for judgment of numerical inequality. *Nature*, 215, 1519-1520. doi:[10.1038/2151519a0](https://doi.org/10.1038/2151519a0)

- Negen, J., & Sarnecka, B. (2015). Is there really a link between exact-number knowledge and approximate number system acuity in young children? *British Journal of Developmental Psychology*, 33, 92-105. doi:[10.1111/bjdp.12071](https://doi.org/10.1111/bjdp.12071)
- Odic, D., Libertus, M., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of approximating area and approximating number. *Developmental Psychology*, 49, 1103-1112. doi:[10.1037/a0029472](https://doi.org/10.1037/a0029472)
- Odic, D., Pietroski, P., Hunter, T., Lidz, J., & Halberda, J. (2013). Young children's understanding of "more" and discrimination of number and surface area. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 451-461. doi:[10.1037/a0028874](https://doi.org/10.1037/a0028874)
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3, 179-197. doi:[10.1163/156856888X00122](https://doi.org/10.1163/156856888X00122)
- Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, 19, 607-614. doi:[10.1111/j.1467-9280.2008.02130.x](https://doi.org/10.1111/j.1467-9280.2008.02130.x)
- Rosenberg, R. D., & Feigenson, L. (2013). Infants hierarchically organize memory representations. *Developmental Science*, 16, 610-621. doi:[10.1111/desc.12055](https://doi.org/10.1111/desc.12055)
- Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development*, 74, 1807-1822. doi:[10.1046/j.1467-8624.2003.00639.x](https://doi.org/10.1046/j.1467-8624.2003.00639.x)
- Rousselle, L., & Noël, M. P. (2008). The development of automatic numerosity processing in preschoolers: Evidence for numerosity-perceptual interference. *Developmental Psychology*, 44, 544-560. doi:[10.1037/0012-1649.44.2.544](https://doi.org/10.1037/0012-1649.44.2.544)
- Rousselle, L., Palmers, E., & Noël, M. P. (2004). Magnitude comparison in preschoolers: What counts? Influence of perceptual variables. *Journal of Experimental Child Psychology*, 87, 57-84. doi:[10.1016/j.jecp.2003.10.005](https://doi.org/10.1016/j.jecp.2003.10.005)
- Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, 108, 662-674. doi:[10.1016/j.cognition.2008.05.007](https://doi.org/10.1016/j.cognition.2008.05.007)
- Shapiro, S. (2000). *Thinking about mathematics: The philosophy of mathematics*. Oxford, United Kingdom: Oxford University Press.
- Shusterman, A., Slusser, E., Halberda, J., & Odic, D. (2016). Acquisition of the cardinal principle coincides with improvement in the acuity of the Approximate Number System in preschoolers. *PLoS One*, 11, Article e0153072. doi:[10.1371/journal.pone.0153072](https://doi.org/10.1371/journal.pone.0153072)
- Starr, A., Libertus, M. E., & Brannon, E. M. (2013a). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences*, 110, 18116-18120. doi:[10.1073/pnas.1302751110](https://doi.org/10.1073/pnas.1302751110)
- Starr, A., Libertus, M. E., & Brannon, E. M. (2013b). Infants show ratio-dependent number discrimination regardless of set size. *Infancy*, 18, 927-941. doi:[10.1111/infa.12008](https://doi.org/10.1111/infa.12008)
- Suanda, S. H., Tompson, W., & Brannon, E. (2008). Changes in the ability to detect ordinal numerical relationships between 9 and 11 months of age. *Infancy*, 13, 308-337. doi:[10.1080/15250000802188800](https://doi.org/10.1080/15250000802188800)
- Szűcs, D., Nobes, A., Devine, A., Gabriel, F. C., & Gebuis, T. (2013). Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children. *Frontiers in Psychology*, 4, Article 444. doi:[10.3389/fpsyg.2013.00444](https://doi.org/10.3389/fpsyg.2013.00444)

- Trick, L., Enns, J. T., & Brodeur, D. A. (1996). Life span changes in visual enumeration: The number discrimination task. *Developmental Psychology, 32*, 925-932. doi:[10.1037/0012-1649.32.5.925](https://doi.org/10.1037/0012-1649.32.5.925)
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101*, 80-102. doi:[10.1037/0033-295X.101.1.80](https://doi.org/10.1037/0033-295X.101.1.80)
- Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1999). What representations might underlie infant numerical knowledge? *Cognitive Development, 14*, 1-36. doi:[10.1016/S0885-2014\(99\)80016-1](https://doi.org/10.1016/S0885-2014(99)80016-1)
- Van Herwegen, J., Ansari, D., Xu, F., & Karmiloff-Smith, A. (2008). Small and large number processing in infants and toddlers with Williams syndrome. *Developmental Science, 11*(5), 637-643. doi:[10.1111/j.1467-7687.2008.00711.x](https://doi.org/10.1111/j.1467-7687.2008.00711.x)
- Wagner, J. B., & Johnson, S. C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition, 119*, 10-22. doi:[10.1016/j.cognition.2010.11.014](https://doi.org/10.1016/j.cognition.2010.11.014)
- Wood, J. N., & Spelke, E. S. (2005). Infants' enumeration of actions: Numerical discrimination and its signature limits. *Developmental Science, 8*(2), 173-181. doi:[10.1111/j.1467-7687.2005.00404.x](https://doi.org/10.1111/j.1467-7687.2005.00404.x)
- Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*, 155-193. doi:[10.1016/0010-0277\(90\)90003-3](https://doi.org/10.1016/0010-0277(90)90003-3)
- Wynn, K. (1992). Child's acquisition of the number words and the counting system. *Cognitive Psychology, 24*, 220-251. doi:[10.1016/0010-0285\(92\)90008-P](https://doi.org/10.1016/0010-0285(92)90008-P)
- Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition, 89*, B15-B25. doi:[10.1016/S0010-0277\(03\)00050-7](https://doi.org/10.1016/S0010-0277(03)00050-7)
- Xu, F., & Arriaga, R. I. (2007). Number discrimination in 10-month-old infants. *British Journal of Developmental Psychology, 25*, 103-108. doi:[10.1348/026151005X90704](https://doi.org/10.1348/026151005X90704)
- Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*, B1-B11. doi:[10.1016/S0010-0277\(99\)00066-9](https://doi.org/10.1016/S0010-0277(99)00066-9)
- Xu, F., Spelke, E., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8*, 88-101. doi:[10.1111/j.1467-7687.2005.00395.x](https://doi.org/10.1111/j.1467-7687.2005.00395.x)
- Zosh, J. M., Halberda, J., & Feigenson, L. (2011). Memory for multiple visual ensembles in infancy. *Journal of Experimental Psychology: General, 140*, 141-158. doi:[10.1037/a0022925](https://doi.org/10.1037/a0022925)