Received: 24 January 2022 | Revised: 1 July 2022 | Accepted: 6 September 2022

DOI: 10.1111/desc.13335

Developmental Science WILEY

RESEARCH ARTICLE

Young children interpret number gestures differently than nonsymbolic sets

Dominic J. Gibson[1,2] | **Talia Berkowitz**[1] | **Jacob Butts**[1] | **Susan Goldin-Meadow**[1] | **Susan C. Levine**[1]

[1]University of Chicago, Chicago, Illinois, USA

[2]foundry10, Seattle, Washington, USA

Correspondence
Dominic J. Gibson, foundry10, 4545 Roosevelt Way, Suite #500, Seattle, WA 98105, USA.
Email: dominic.gibson@gmail.com

https://youtu.be/WtVziFN1yuI

Funding information
the Hymen Milgrom Supporting Organization; NCER, Grant/Award Number: R305B090025; National Science Foundation Graduate Research Fellowship, Grant/Award Number: DGE-1144082

Abstract

Researchers have long been interested in the origins of humans' understanding of symbolic number, focusing primarily on how children learn the meanings of number words (e.g., "one", "two", etc.). However, recent evidence indicates that children learn the meanings of number gestures before learning number words. In the present set of experiments, we ask whether children's early knowledge of number gestures resembles their knowledge of nonsymbolic number. In four experiments, we show that preschool children ($n = 139$ in total; age $M = 4.14$ years, $SD = 0.71$, range = 2.75–6.20) do not view number gestures in the same the way that they view nonsymbolic representations of quantity (i.e., arrays of shapes), which opens the door for the possibility that young children view number gestures as symbolic, as adults and older children do.

KEYWORDS
cognitive development, gesture, number, symbols

Highlights

- Children were more accurate when enumerating briefly-presented number gestures than arrays of shapes, with a shallower decline in accuracy as quantities increased.
- We replicated this finding with arrays of shapes that were organized into neat, dice-like configurations (compared to the random configurations used in Experiment 1).
- The advantage in enumerating briefly-presented number gestures was evident before children had learned the cardinal principle.
- When gestures were digitally altered to pit handshape configuration against number of fingers extended, children overwhelmingly based their responses on handshape configuration.

1 | INTRODUCTION

Humans are born able to reason about nonsymbolic quantities—even infants can keep track of precisely 1–3 objects and make approximate judgments about larger quantities (Feigenson et al., 2004). However, everything from landing a Mars rover to filing taxes requires an understanding of symbolic number (i.e., representing specific quantities with agreed upon symbols like "three", "17", etc.). Accordingly, disparities across children in their grasp of symbolic number, which emerge by kindergarten, predict later success in school, the workplace, and everyday life (Duncan et al., 2007; Geary et al., 2018). As a result, the transition from thinking about quantities non-symbolically to also thinking about them symbolically is an important milestone, both historically and in the lives of young children.

Although people communicate about number in several ways (numerals, words, tallies, gestures, and so forth), most investigations into children's acquisition of symbolic number focus on number words (i.e., "one", "two", "three", etc.; Sarnecka & Carey, 2008; Sarnecka & Lee, 2009; Wynn, 1990, 1992). These studies show that learning the meanings of number words does not come easily to children. Upon learning to count around the age of two, children initially lack an understanding of what each number word means (i.e., which words refer to which quantities). They slowly learn the meanings of each of the first few number words, one at a time, before finally, at around age four, grasping the cardinal principle––understanding that the last number reached when counting a set of objects represents the value of that set (Wynn, 1990, 1992).

A body of research suggests that gestures may play an important supportive role in children's math learning (see Goldin-Meadow et al., 2014, for a review and Crollen et al., 2011, for a contrasting view). For instance, many have noticed that children frequently point to items while counting and have argued that this helps children implement the counting principles correctly (e.g., Alibali & DiRusso, 1999; Fuson, 1988; Gelman & Gallistel, 1978; Gordon et al., 2019). Less research has focused on the role that children's representational or cardinal number gestures (e.g., holding up three fingers to indicate "three") play in number learning. The exceptions are a study by Gibson et al. (2019), who found that children's use of these gestures in concert with speech predicted the likelihood of learning from a cardinal number intervention, and another study by Orrantia et al. (2022), who found that incorporating these gestures into an intervention involving counting and labeling sets improved children's cardinal number understanding.

Why might representational or cardinal number gestures (from here on referred to just as "number gestures") play a supportive role in children's verbal number learning? One possibility is that number gestures share commonalities with both nonsymbolic quantities (e.g., sets of objects) and symbolic number (e.g., number words; Di Luca & Pesenti, 2011; Wiese, 2007). Like nonsymbolic quantities, number gestures are iconic, item-based representations of number (the form of a number gesture reflects its meaning––the number of fingers displayed in a gesture corresponds exactly to the number of items the gesture is meant to describe). This may make number gestures easier to learn than number words, whose forms are arbitrarily related to quantity (for reviews on iconicity's role in symbol acquisition, see Dingemanse et al., 2015; Perniss & Vigliocco, 2014). On the other hand, like number words, number gestures can also be used as summary symbols (i.e., a single number gesture can be used to represent a specific quantity of items), which may help children associate them with number words. By virtue of sharing these similarities with both nonsymbolic quantities and symbolic number, number gestures could be a useful bridge between nonsymbolic and symbolic number (Gunderson et al., 2015; Ifrah, 2000, Orrantia et al., 2022). However, the extent to which children appreciate both the iconic properties, and the symbolic properties, of number gestures is an open question.

With respect to children's appreciation of the iconicity of number gestures, we know that children can accurately use number gestures to label the number of items in a set before they are able to do so using corresponding number words (Gibson, Butts, Goldin-Meadow, Levine, in prep; Gunderson et al., 2015). Specifically, Gunderson et al. (2015) presented children with pictures of varying sets of items and asked children to label them, using either number words or number gestures. Children who were still in the process of learning the basic meanings of number words were more accurate when labeling the sets using number gestures than number words, particularly for set sizes that were immediately above their current level of verbal number knowledge. In contrast, cardinal principle knowers were accurate when labeling sets with either number words or number gestures.

However, Nicoladis et al. (2018) have called into question whether Gunderson et al.'s (2015) findings are truly due to children appreciating the iconicity of number gestures. They argue, instead, that children initially learn number gestures as arbitrary symbols and only later come to appreciate the correspondence between fingers and items in a set. Their evidence for this view is that children are more accurate when matching quantities to canonical number gestures (e.g., raising index and middle finger to indicate two) than non-canonical number gestures (e.g., raising the index finger on each hand to indicate two). Despite the advantage of canonical forms, this task could still be solved by iconicity since both canonical and non-canonical forms are iconic, and the advantage of canonical forms could reflect the fact that they are more familiar to children than non-canonical forms. The wider literature on iconicity does not resolve this question since it suggests children can appreciate gesture's iconicity by at least 26 months (prior to when most children comprehend numbers) but are not so dependent on iconicity when learning new words or gestures that we would expect them to necessarily pick up on the iconicity of every gesture (e.g., Namy, 2008). Moreover, previous studies of iconic gestures generally involve gestures that are nearly identical to the action they are referencing, whereas the iconicity of number gestures might be less obvious. As a result, even though Gunderson et al.'s (2015) results suggest that children do learn to associate specific quantities with number gestures before they make analogous mappings with number words, we do not yet know whether they do so because of the iconicity of number gesture.

Nor do we know whether early number gestures refer to sets, as symbolic number words do. One interpretation of Gunderson et al. (2015) is that children view number gestures as item-based representations as they would any other set of objects (e.g., "finger, finger, finger"), rather than as summary symbols representing a set (e.g. "three"). On this account, children may map number gestures to sets of objects in the same way that they have been shown to match two sets of objects based on set size (Huttenlocher et al., 1994; Mix, 2008). This differs from symbolic number, in which a single symbol, like a number word, is mapped to a specific quantity. This is a defining and critical feature of symbolic number since it is what enables us to represent very large quantities without an equally large number of tallies.

Therefore, in the current study, we asked whether children really view number gestures purely as item-based representations (e.g., collections of dots) or whether children can view them as whole-gestalts. If the latter is true, it could suggest that when children successfully map a number gesture to a corresponding quantity (as shown by Gunderson et al., 2015), this is best interpreted as a single symbol mapped to a specific quantity of items, and thus, number gestures may represent an early form of symbolic number knowledge.

We tested the hypothesis that young children view number gestures differently from other nonsymbolic sets of objects in four experiments, three involving children who recently learned the meanings of number words (Experiments 1, 2, and 4), and one involving children who were still in the process of learning the meanings of number words (Experiment 3). In Experiments 1–3, we used as a benchmark a characteristic of children's representation of nonsymbolic quantities that has been observed in many studies of infants, adults, and nonhuman animals. Specifically, when prevented from counting to attain the number of items in a set (e.g., if a set of shapes is flashed very quickly on a screen), children (and adults) are forced to estimate using the approximate number system (ANS), resulting in decreasing accuracy and increasing variance in children's estimates as the number of items to be enumerated increases (e.g. Le Corre & Carey, 2007; Moyer & Landauer, 1967; Whalen et al., 1999). As set sizes increase, accuracy continuously falls off, with the exception that number in very small sets of items (1–3) is accurate (because these sets are subitizable, humans can immediately perceive exactly how many items there are in the display, (e.g., Kaufman et al., 1949; Mandler & Shebo, 1982). Beyond three items, children, as well as adults, need to count in order to consistently arrive at an exact answer, and the likelihood of arriving at an exact answer through estimation decreases as a function of the size of the set.

Here we ask whether the same pattern arises for gesture––do we observe steeply decreasing accuracy beyond the subitizable range when children are tasked with quickly enumerating the number of fingers raised in a number gesture? If children conceive of number gestures as nonsymbolic sets of items, they should struggle to arrive at the correct answer when asked to enumerate sets beyond three. However, if children view conventional number gestures in the same way that numerate adults view numerals and number words, then they should be able to arrive at the correct answer relatively quickly for sets beyond three, resulting in a relatively small decrease in accuracy as numbers increase.[1]

In Experiment 1, we measured children's ability to enumerate the number of fingers raised in a number gesture and, for comparison, the number of shapes in an array in a speeded task that prevented counting. In Experiment 2, we added a second version of the shapes task in which shapes were arranged in dice configurations to control for the greater visual organization of gestures compared to the random dot arrays used in Experiment 1. We began our investigation with children who had already mastered the cardinal principle (i.e., children who had learned the basic meanings of the number words in their count list) since previous research suggests that fast enumeration tasks are most appropriate for children who have already learned the cardinal principle (e.g., Le Corre & Carey, 2007). However, in Experiment 3, we extended our investigation to children who had not yet learned the cardinal principle to test whether children's fast enumeration of number gestures differs from their fast enumeration of shapes even before they fully understand the basic meanings of the number words in their count list. Finally, in Experiment 4, we digitally altered number gestures so that the handshape *gestalt* conflicted with the actual *number* of fingers raised, as a strong test of the hypothesis that children attend to the overall gestalt of number gestures, rather than the number of individual fingers, when making these judgments.

2 | EXPERIMENT 1: ENUMERATION OF GESTURES VERSUS SHAPES

2.1 | Methods

2.1.1 | Participants

Thirty-three children (19 female) participated in the first experiment. The mean age was 4.59 years (*SD* = 0.76, range = 3.35–6.21 years). Participants came from a range of socioeconomic backgrounds (see Table 1 for full demographics for each experiment). Family income was assessed in terms of categories (Less than $15,000; $15,000 to $34,999; $35,000 to $49,999; $50,000 to $74,999; $75,000 to $99,999; and $100,000 or more) and ranged from less than $15,000 per year to more than $100,000 with average family income falling in the $50,000 to $75,000 per year category.

2.1.2 | Procedure

Participants were selected for the current study from a larger sample of preschool students who completed a number battery, before being sorted into separate studies based on their knower-level. All participants were first asked to count as high as they could and then to complete the Give-a-Number task (Wynn, 1990, 1992). Only children who successfully demonstrated knowledge of the cardinal principle were included in Experiment 1.[2]

Participants in all four experiments were either recruited through a database of interested families and tested in the lab or recruited through urban preschools and daycares and tested in a quiet corner or room of their preschool or daycare. Children were included only if their parents completed and returned a consent form that was either sent home or completed on site for children tested in the lab. Unless otherwise noted (Experiments 2 and 4), no participants were dropped from the study.

2.1.3 | Give-a-Number

The Give-a-Number task was used to determine each child's knower level, which specifies the highest number word for which a child

TABLE 1 Demographics by experiment

	Experiment 1 ($n = 33$)	Experiment 2 ($n = 51$)	Experiment 3 ($n = 21$)	Experiment 4 ($n = 34$)
Gender, %				
Boys	42 (14/33)	49 (25/51)	57 (12/21)	53 (18/34)
Girls	58 (19/33)	51 (26/51)	43 (9/21)	44 (15/34)
Missing				3 (1/34)
Race, %				
Asian		8 (4/51)		
Black	27 (9/33)	20 (10/51)	33 (7/21)	9 (3/34)
Native Hawaiian or other Pacific Islander		2 (1/51)		
Multiracial/other	6 (2/33)	2 (1/51)	19 (4/21)	9 (3/34)
White	45 (15/33)	55 (28/51)	5 (1/21)	41 (14/34)
Missing	21 (7/33)	14 (7/51)	43 (9/21)	41 (14/34)
Ethnicity, %				
Hispanic/Latino	3 (1/33)	8 (4/51)	29 (6/21)	18 (6/34)
Not Hispanic/Latino	76 (25/33)	76 (39/51)	33 (7/21)	47 (16/34)
Missing	21 (7/33)	16 (8/51)	38 (8/21)	35 (12/34)
Household income, %				
<$15,000	6 (2/33	6 (3/51)	14 (3/21)	6 (2/34)
$15,000–$34,999	9 (3/33	6 (3/51)	24 (5/21)	9 (3/34)
$35,000–$49,999	6 (2/33	4 (2/51)	5 (1/21)	
$50,000–$74,999	3 (1/33	10 (5/51)		3 (1/34)
$75,000–$99,999	6 (2/33	4 (2/51)		6 (2/34)
>$100,000	48 (1/33	55 (28/51)	5 (1/21)	35 (12/34)
Missing	21 (7/33	16 (8/51)	52 (11/21)	41 (14/34)
Parent education, %				
Less than high school			10 (2/21)	6 (2/34)
High school or GED	6 (2/33)		19 (4/21)	9 (3/34)
At least 1 year of college	15 (5/33)	4 (2/51)	10 (2/21)	3 (1/34)
Associate/2 years degree	3 (1/33)	8 (4/51)		
Bachelor's/4 years degree	18 (6/33)	31 (16/51)	5 (1/21)	15 (5/34)
Some graduate		4 (2/51)	5 (1/21)	
Graduate degree	36 (12/33)	41 (21/51)	5 (1/21)	29 (10/34)
Missing	21 (7/33)	12 (6/51)	48 (10/21)	38 (13/34)

understands the cardinal value (Wynn, 1990, 1992). Children were presented with 15 plastic fish and asked to place a certain number of fish into a clear plastic bowl (called "the pond"). If a child gave the wrong number of fish, the experimenter gave the child an opportunity to correct the mistake by saying, "But I asked for N fish! Let's check. [Experimenter and child count fish.] Can you put N fish in the pond?" Children's final answers were recorded. The experimenter always began by asking the child to place one fish in the pond. The experimenter then increased the number requested by one fish every time the child answered correctly, and decreased the number requested by one fish every time the child answered incorrectly, as in Wynn (1990). Children were classified as an N knower when N was the highest number for which they responded correctly on two out of three requests for N fish and gave the experimenter N fish less than half as often when asked for more than N fish than when asked for N fish. If children succeeded on all numbers up to 6, they were considered cardinal principle knowers. As noted earlier, only children who successfully demonstrated knowledge of the cardinal principle were included in Experiment 1.

2.1.4 | Fast Shapes

The Fast Shapes task consisted of four blocks of seven trials (28 trials total) presented in a fixed, random order. In each block, children saw triangles, squares, diamonds, or plus-signs on a laptop computer

Developmental Science WILEY

screen. First, children were given instructions and familiarized with the task. They were told that some shapes were going to flash up on the screen and that it would happen too quickly to count; their task was to try to guess how many shapes they saw. We then familiarized them to the displays by showing them sets of 1–15 circles, labeling each set along with the child as the sets appeared. After this, the experimenter asked if the child was ready and, if so, began the task. On each test trial, the dot array appeared on the screen for 1 s before disappearing. The experimenter asked each child to say how many shapes they thought they saw. If the child was reluctant to answer, the experimenter prompted the child to "just make your best guess." If the child tried counting during the brief exposure, the experimenter reminded the child, "Remember it goes too fast to count, so just try to make your best guess––how many shapes do you think you saw?" If the experimenter noticed that the child had looked away and missed the trial, the experimenter repeated the trial and recorded that the trial had been shown twice. If children changed their response, only the final response was included in the analyses. Children were tested on seven set sizes (1, 2, 3, 4, 5, 10, and 14 shapes), which repeated in each block. We calculated children's accuracy for each set size as the number of trials they got exactly right, divided by the total number of trials to which they gave a numeric response.

2.1.5 | Fast Gestures

The Fast Gestures task mirrored the Fast Shapes task. Again, the task consisted of four blocks of seven trials (28 trials total) presented in a fixed, random order. In each block, children saw a different person's hands with various numbers of fingers extended. The same seven set sizes were shown in each block as in the Fast Shapes task (1, 2, 3, 4, 5, 10, and 14). For sets less than or equal to five, only one hand appeared; for the 10-finger trial, two hands appeared; and for the 14-finger trial, three hands appeared. All gestures were canonical American number gestures or combinations of canonical gestures (i.e., 14 was made up of three hands displayed in a row: 5-, 5-, and 4-fingered gestures). Children were given the same instructions as in the Fast Shapes task and were shown parallel familiarization displays of gestures displaying 1–15 fingers. In short, the procedure for the Fast Gestures task was identical to the procedure for the Fast Shapes task, except that gestures were used rather than sets of shapes and children were asked to guess how many fingers they saw.

2.2 | Results

We predicted greater accuracy for number gestures than for sets of shapes, particularly on trials above the subitizable range (i.e., more than three). To test this hypothesis, we conducted a two-way repeated-measures ANOVA on accuracy with the Greenhouse–Geisser correction, with the within-subjects factors of task (Fast Gestures vs. Fast Shapes) and set size (1, 2, 3, 4, 5, 10, and 14). We found a main effect of task ($F(1,32) = 66.41$, $p < 0.001$, $\eta_p^2 = 0.676$), a main effect of set size

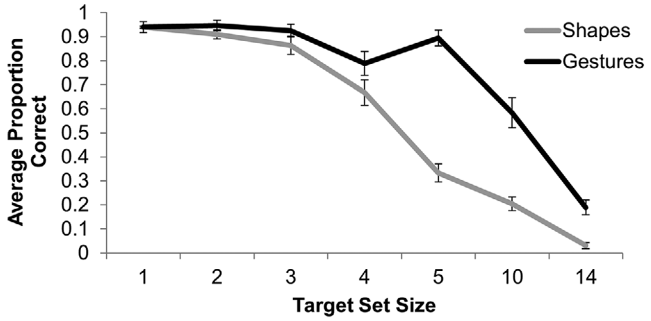

FIGURE 1 Accuracy on Fast Shapes and Fast Gestures tasks by set size. Error bars represent the standard error in accuracy for each set size across children

TABLE 2 Predictors of accuracy for set sizes 1 to 5

	b	S.E.	df	*t*	Sig.
Intercept	0.59	0.04	173.13	14.60	<0.001
Task	0.16	0.02	314.64	7.36	<0.001
Set size	−0.27	0.02	310.60	−11.77	<0.001
Task*set size	0.12	0.01	310.56	8.42	<0.001

($F(4.1, 129.8) = 150.40$, $p < 0.001$, $\eta_p^2 = 0.825$), and a significant interaction of task and set size ($F(3.8, 123) = 14.66$, $p < 0.001$, $\eta_p^2 = 0.314$). Accuracy by task and set size is depicted in Figure 1. The raw counts of children's responses to each set size are available in the Appendix (Appendix Figure A1).

Whenever children's performance differed by task, children were more accurate on the Fast Gestures task than the Fast Shapes task. The greatest differences between the Fast Gestures and Fast Shapes task were seen on set sizes **5** ($t(32) = 10.29$, $p < 0.001$, $d = 1.79$) and **10** ($t(33) = 5.02$, $p < 0.001$, $d = 0.86$). There were significant but more moderate differences in accuracy between the tasks on set sizes **4** ($t(32) = 2.19$, $p = 0.036$, $d = 0.38$) and **14** ($t(32) = 10.29$, $p < 0.001$, $d = 0.44$). There were no significant differences in accuracy on set sizes **1** ($t(32) = 0$, $p < 1.000$, $d = 0$), **2** ($t(32) = 1.15$, $p = 0.257$, $d = 0.20$), or **3** ($t(32) = 1.14$, $p = 0.264$, $d = 0.20$); children performed close to ceiling on both tasks for these values, which are within the subitizable range.

We also predicted that, within the one-handed gestures (1–5), there would not be the same decline in accuracy on the Fast Gestures Task beyond the subitizable range that is typically shown in the Fast Shapes task (e.g., Le Corre & Carey, 2007). To test this prediction, we used a linear mixed model with subject as a random effect and Task (Shapes vs. Gestures) and Set Size (1–5; continuous) as fixed effects. We centered Set Size at 3 in order to get an estimate of the main effect of task at the top of the subitizable range. The intercept was allowed to vary randomly. The results of this estimation are presented in Table 2.

Results indicate that there was a main effect of Task, whereby children were more accurate in the Fast Gestures task than the Fast Shapes task. There was also a main effect of Set Size, whereby accuracy decreased as set size increased. Most importantly, there was a significant interaction between Task and Set Size. The positive coefficient of the interaction term indicates that accuracy did not fall as steeply between set sizes 1 to 5 in the Fast Gestures task as in the Fast Shapes task, indicating the predicted gesture advantage for set sizes beyond the subitizable range.

The results of these analyses show that children were more accurate and displayed a more attenuated decrease in performance as set size increased in the Fast Gestures task than in the Fast Shapes task. A visualization of the distribution of children's responses for each set size on each task can be found in the Appendix (Appendix Figure A4).

2.3 | Experiment 1 discussion

Replicating previous studies, accuracy on the Fast Shapes task declined steadily beyond set sizes 1–3, reflecting children's reliance on the ANS to roughly judge the number of items in larger sets. In contrast, there was a shallower decline in accuracy on the Fast Gestures task across set sizes 1 to 5 and children were more accurate for quantities above the subitizable range, compared to the Fast Shapes task. This pattern may indicate that, on the Fast Gestures task, children did not need to rely on estimates using their ANS and instead recognized each number gesture in the same holistic way that an adult would recognize a number word or numeral.

An alternative possibility is that children relied on the same mechanisms in the two tasks, but that the spatial organization of number gestures made them easier for young children to enumerate, compared to the arrays of shapes used in this task. Conventional number gestures naturally display fingers in more organized arrays than the random configurations of shapes displayed in our Fast Shapes task. Moreover, although we used hands from different individuals to form the gestures in each block, the basic configuration of the gestures was the same within each set size (e.g., index and middle finger were always extended in the "two" gesture). In the dot arrays, however, set sizes were presented in different configurations across trials with no specific organization. Therefore, to test this alternative explanation, in Experiment 2 we examined whether the organization of the arrays could explain the difference between Fast Gestures and Fast Shapes.

3 | EXPERIMENT 2: ENUMERATION OF SHAPES VERSUS DICE VERSUS GESTURES

To test the alternative explanation of the results of Experiment 1, we replicated the experiment with a new group of participants, enlarged the sample, and added a third task––Fast Dice. The Fast Dice task was identical to the other two tasks except the shapes were organized into canonical dice configurations of up to five shapes (e.g., 10 was represented by two identical sets of five shapes, comparable to the two-handed gesture for 10). Thus, the Fast Dice task was designed to more closely match several potentially important visual characteristics of the Fast Gestures task. In addition, as in the Fast Gestures task, the Fast Dice task presented each set size using the same configuration of shapes across the four trials of that set size. If the greater accuracy in the original Fast Gestures task was due to these features, and to the more organized configurations of gestures compared to random dot arrays, then children should also be more accurate when estimating the number of shapes in a dice array than the number of shapes in a random array (Fast Shapes task). Conversely, if the gesture advantage observed in Experiment 1 is due to children's interpreting gestures as they do words and not merely to gesture's stable spatial organization, then the gesture advantage should extend to Experiment 2. Specifically, we predicted that children's enumeration of gestures would be more accurate than their enumerations not only of sets of randomly arrayed shapes (as observed in Experiment 1), but also of organized arrays (i.e., dice configurations). This result would support the hypothesis that the advantage of gestural representations of sets was not due solely to the organized nature of the input, but also to children's holistic interpretation of these gestures.

3.1 | Methods

3.1.1 | Participants

Fifty-one children (26 female) participated in the study. One additional participant was dropped because they completed only one of the three enumeration tasks. Two other participants completed two of the three enumeration tasks (one completed Fast Gestures and Fast Dice; the other completed Fast Gestures and Fast Shapes) and therefore were not included in analyses that compared performance on all three tasks. The mean age of the participants was 4.03 years (*SD* = 0.57, range = 2.75–5.06 years). Average family earnings fell in the \$50,000 and \$75,000 per year category.

3.1.2 | Procedure

The procedure was identical to Experiment 1 except for the addition of the Fast Dice task. Participants received the Give-N task as described in Experiment 1 to determine their knower level. If they were cardinal principle-knowers, they were included in the study and completed three enumeration tasks: Fast Shapes (see Experiment 1), Fast Gestures (see Experiment 1) and Fast Dice. Each participant received all three tasks, and the order of the tasks was counterbalanced across participants.

3.1.3 | Fast Dice

The Fast Dice task closely mirrored the other two enumeration tasks. Again, the task consisted of four blocks of seven trials (28 trials total) presented in a fixed, random order. In each block, children saw arrays made up of different basic shapes. The same seven set sizes used in the other enumeration tasks were shown in each block (1, 2, 3, 4, 5, 10, and 14). For sets less than or equal to five, the shapes were organized into configurations like those used on traditional six-sided dice. For set size

10, two groups of five shapes (each with the same configuration as the set size 5 trials) were displayed. For set size 14, two groups of 5 and one group of 4 were displayed. The procedure and instructions were the same as those used in the Fast Shapes task (see Experiment 1).

3.2 | Results

3.2.1 | Replication of Experiment 1

As a direct replication of our results from Experiment 1, we ran a 2 (Task: Fast Shapes vs. Fast Gesture) by 7 (Set Size: 1, 2, 3, 4, 5, 10, and 14) repeated measures ANOVA on accuracy with the Greenhouse–Geisser correction. We found a main effect of task ($F(1.0, 49.0) = 65.63$, $p < 0.001$, $\eta_p^2 = 0.573$), a main effect of set size ($F(4.0, 198.1) = 225.60$, $p < 0.001$, $\eta_p^2 = 0.822$), and a significant interaction of task and set size ($F(2.2, 192.6) = 29.48$, $p < 0.001$, $\eta_p^2 = 0.423$). Participants were more accurate on the Fast Gestures task than the Fast Shapes task for set sizes **4** ($t(49) = 3.47$, $p = 0.001$, $d = 0.49$), set size **5** ($t(49) = 14.84$, $p < 0.001$, $d = 2.10$), and set size **10** ($t(49) = 4.31$, $p < 0.001$, $d = 0.61$). There was a small difference in accuracy between the tasks on set size **3** ($t(49) = 2.01$, $p = 0.050$, $d = 0.28$). There were no significant differences in accuracy on set sizes **1** ($t(49) = -1.74$, $p < 0.088$, $d = 0.25$), **2** ($t(49) = 0.57$, $p = 0.569$, $d = 0.08$), or **14** ($t(49) = 1.75$, $p = 0.086$, $d = 0.25$).

3.2.2 | Fast Shapes, Fast Dice, versus Fast Gestures

We conducted a repeated-measures ANOVA on accuracy with the Greenhouse–Geisser correction, with the within-subjects factors of task (Fast Gestures vs. Fast Shapes vs. Fast Dice) and set size (1, 2, 3, 4, 5, 10, and 14). We found a main effect of task ($F(1.4, 69.2) = 52.86$, $p < 0.001$, $\eta_p^2 = 0.524$), a main effect of set size ($F(3.7, 178.4) = 274.50$, $p < 0.001$, $\eta_p^2 = 0.851$), and a significant interaction of task and set size ($F(6.1, 293.4) = 25.48$, $p < 0.001$, $\eta_p^2 = 0.347$). Accuracy by task and set size is depicted in Figure 2. The raw counts of children's responses to each set size are available in the Appendix (Appendix Figure A2).

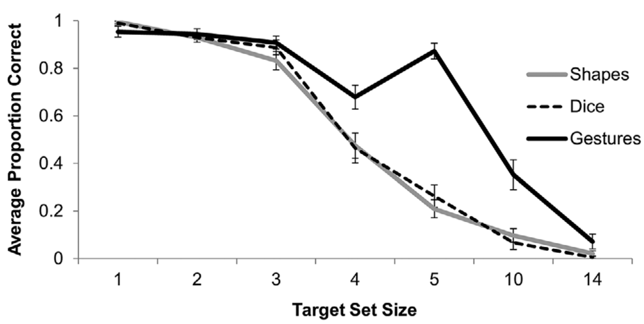

FIGURE 2 Accuracy on Fast Shapes, Fast Dice, and Fast Gestures tasks by set size. Error bars represent the standard error in accuracy for each set size across children

TABLE 3 Predictors of accuracy for set sizes 1 to 5 (Dice vs. Gestures)

	b	S.E.	df	*t*	Sig.
Intercept	0.55	0.04	392.57	14.62	<0.001
Task	0.16	0.02	454.41	7.55	<0.001
Set size	−0.34	0.02	451.17	−14.18	<0.001
Task*set size	0.15	0.02	451.17	9.85	<0.001

To better understand these effects, we also ran separate ANOVAs comparing the two shapes tasks (Fast Shapes and Fast Dice) and comparing the Fast Dice task to the Fast Gestures Task. First, we compared accuracy on the two types of Fast Shapes tasks by running a 2 (Task: Fast Shapes vs. Fast Dice) by 7 (Set Size: 1, 2, 3, 4, 5, 10, and 14) repeated-measures ANOVA on accuracy with the Greenhouse–Geisser correction. We found a main effect of set size ($F(3,143) = 240.85$, $p < 0.001$, $\eta_p^2 = 0.834$), but no effect of task ($F(1,48) = 0.31$, $p = 0.583$, $\eta_p^2 = 0.006$), and no task by set size interaction ($F(4.1,199) = 1.24$, $p = 0.293$, $\eta_p^2 = 0.025$). Thus, we found no evidence that children benefited from the more organized configuration of the shapes in the Fast Dice task, compared to the Fast Shapes task.

Next, we ran a 2 (Task: Fast Dice vs. Fast Gesture) by 7 (Set Size: 1, 2, 3, 4, 5, 10, and 14) repeated measures ANOVA on accuracy with the Greenhouse–Geisser correction. We found a main effect of task ($F(1.0, 49.0) = 56.98$, $p < 0.001$, $\eta_p^2 = 0.538$), a main effect of set size ($F(3.7, 181.7) = 229.11$, $p < 0.001$, $\eta_p^2 = 0.824$), and a significant interaction of task and set size ($F(3.6, 176.3) = 29.48$, $p < 0.001$, $\eta_p^2 = 0.376$). Participants were more accurate on the Fast Gestures task than the Fast Dice task for set sizes **4** ($t(49) = 2.78$, $p = 0.008$, $d = 0.39$), **5** ($t(49) = 12.57$, $p < 0.001$, $d = 1.78$), and **10** ($t(49) = 5.33$, $p < 0.001$, $d = 0.75$). There was also a small but significant difference in accuracy between the tasks on set size **14** ($t(49) = 2.10$, $p = 0.041$, $d = 0.30$). There were no significant differences in accuracy on set sizes **1** ($t(49) = -1.55$, $p = 0.128$, $d = 0.22$), **2** ($t(49) = 0.39$, $p = 0.7$, $d = 0.05$), or **3** ($t(49) = 0.45$, $p = 0.652$, $d = 0.06$).

We also repeated our analyses from Experiment 1, this time comparing Fast Gestures and Fast Dice using a linear mixed model with subject as a random effect and Task (Dice vs. Gestures) and Set Size (1–5; continuous) as fixed effects (Table 3).

The results closely mirrored those of Experiment 1, showing a main effect of Task and Set Size, reflecting children's better performance enumerating gestures and smaller sets, as well as a significant interaction between Task and Set Size, reflecting the fact that children's accuracy was not as negatively impacted by increasing set sizes in the Fast Gestures task as on the Fast Dice task.

3.3 | Experiment 2 discussion

We replicated Experiment 1 by showing that children were more accurate at quickly enumerating gestures (Fast Gestures) than randomly arranged shapes (Fast Shapes) for non-subitizable set sizes. There was, however, no significant difference in accuracy between Fast Gestures and Fast Shapes for set size 14. This departure from Experiment 1 is not surprising given the relatively low performance on set size 14 trials in both experiments, and the fact that children rarely, if ever, encounter three-handed number gestures.

Experiment 2 took the phenomenon a step further and asked whether the spatial organization of the arrays could explain the difference in children's accuracy on Fast Gestures and Fast Shapes. We found no evidence to suggest that children benefited from the greater spatial organization of the arrays used in the Fast Dice task, compared to the Fast Shapes task. This result suggests that children's greater accuracy on the Fast Gesture task compared to the Fast Shapes task, as observed in Experiment 1 and replicated in Experiment 2, likely does not stem from the greater spatial organization of the items within the display. Instead, the data provide support for our hypothesis that children view number gestures holistically. Even when *shapes* are presented in neatly organized arrays, there is a steep decline in performance across set sizes 1 through 5, reflecting children's inability to subitize quantities above 3 or 4. In contrast, there is relatively little decline in children's accuracy across *gestures* for 1 through 5. This pattern is reminiscent of the pattern of children's mappings between numerals and number words (two types of arbitrary symbols) (Hurst et al., 2017), suggesting that children who have just recently learned the cardinal principle may already view and recognize number gestures as whole gestalts and not merely item-based representations.

4 | EXPERIMENT 3: SUBSET KNOWERS, DICE VERSUS GESTURES

The results of Experiments 1 and 2 reveal that children who knew the cardinal principle can label number gestures for 1 through 5 with high accuracy without counting. It is therefore possible that children view number gestures holistically only after they have a basic understanding of number words. To address this possibility, in Experiment 3, we tested subset-knowers (children understood only the first four or fewer number words) on the Fast Gestures and Fast Dice tasks. By testing children who have not yet mastered the cardinal meanings of number words, we can assess whether holistic interpretations of number gestures depends on, or predates, understanding the meanings of number words.

4.1 | Methods

4.1.1 | Participants

Twenty-one children (nine female) participated in this study. Their mean age was 3.94 years (*SD* = 0.75, range = 3.52–4.56 years). On average, participants came from a lower socioeconomic background than participants in the prior two experiments, with average family income falling in the $15,000 to $35,000 per year category.

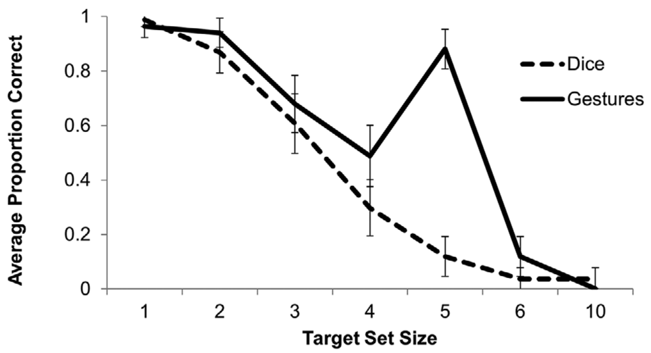

FIGURE 3 Accuracy on Fast Dice and Fast Gestures tasks by set size. Error bars represent the standard error in accuracy for each set size across children

4.1.2 | Procedure

The procedure was identical to Experiments 1 and 2 with a few exceptions. First, participants completed the experiment only if they were subset-knowers rather than cardinal principle knowers, as in the first two experiments. As in the previous experiments, children's knower-levels were determined using the Give-a-Number task. Second, participants received only the Fast Dice and Fast Gestures enumeration tasks, the order of which was counterbalanced across participants. Finally, participants were tested on set sizes 1, 2, 3, 4, 5, 6, and 10. For both the Fast Dice and Fast Gestures tasks, set size 6 was presented as groups of 5 and 1, separated by a space, and set size 10 was presented as two groups of five, to ensure the stimuli were comparable in both tasks.

4.2 | Results

We determined participants' knower-levels using the Give-a-Number task, and found that three participants were one-knowers, five were two-knowers, nine were three-knowers, and 4 were four-knowers. Given our primary goal of understanding how children map number words to number gestures prior to learning the cardinal principle, we did not separate children by knower-level in any of the following analyses.

We conducted a two-way repeated-measures ANOVA on accuracy with the Greenhouse–Geisser correction, with the within-subjects factors of task (Fast Gestures vs. Fast Dice) and set size (1, 2, 3, 4, 5, 6, and 10). We found a main effect of task ($F(1,20) = 38.38$, $p < 0.001$, $\eta_p^2 = 0.657$), a main effect of set size ($F(3.34, 66.85) = 98.75$, $p < 0.001$, $\eta_p^2 = 0.832$), and a significant interaction of task and set size ($F(3.63, 72.55) = 19.55$, $p < 0.001$, $\eta_p^2 = 0.494$). Accuracy by task and set size is depicted in Figure 3. The raw counts of children's responses to each set size are available in the Appendix (Appendix Figure A3).

Overall, accuracy was greater on the Fast Gestures task (Mean = 0.58; *SD* = 0.11) than the Fast Dice task (Mean = 0.42; *SD* = 0.10; $t(20) = 6.20$, $p < 0.001$, $d = 1.35$). Broken down by target set size, the difference between Fast Gestures and Fast Dice was significant only at set size 5 ($t(20) = 11.61$, $p < 0.001$, $d = 2.53$; all other *p*'s > 0.092).

TABLE 4 Predictors of accuracy for set sizes 1 to 5

	b	S.E.	df	*t*	Sig.
Intercept	0.36	0.06	193.67	5.94	<0.001
Task	0.21	0.04	186.00	5.80	<0.001
Set size	–0.40	0.04	186.00	−9.68	<0.001
Task*set size	0.17	0.03	186.00	6.47	<0.001

Although the difference between the two tasks at set size 4 was not significant, the size of the difference was comparable to that found in Experiments 1 and 2 for set size 4. Moreover, participants were still significantly more accurate *overall* on the Fast Gestures than Fast Dice even when set size 5 was removed ($t(20) = 2.19$, $p = 0.041$, $d = 0.48$), suggesting that subset knowers' greater accuracy with gestures was not driven entirely by their performance on set size 5.

As in Experiments 1 and 2, we estimated a linear-mixed model to test whether accuracy declined at the same rate between Set Size 1 to 5 in the two tasks. The results of this analysis are displayed in Table 4. In line with the results of Experiments 1 and 2 on cardinal principle knowers, we found a main effect of task, such that participants were more accurate on the Fast Gestures task than the Fast Shapes task, and a main effect of set size, such that accuracy decreased with set size. Again, there was a significant interaction between format and set size: participants' accuracy in the Fast Gestures task declined less rapidly across set sizes 1 to 5 than their accuracy on the Fast Dice task.

These analyses suggest a similar pattern of results in subset knowers, compared to cardinal principle knowers, with some slight differences. Although there was no significant difference in accuracy for many of the set sizes, the distribution of children's errors provides evidence that they may be solving the Fast Gestures task differently from the Fast Dice task. For instance, when shown 10 fingers (two hands each displaying five fingers, gray dotted line), children never correctly labeled the set "ten"; however, they did often label this display "five". In fact, children labeled set size 10 as "five" on 69% of trials on the Fast Gestures task, compared to only 18% of trials in the Fast Dice task. Likewise, the most common response on Fast Gestures for set size 6, was "one" (27% of trials), likely reflecting the fact that one of the hands was making the gesture for 1. Children also frequently responded with combinations of number words for set size 6 on the Fast Gestures task, such as "five and one". If we include correct combinations as correct responses, then children's average accuracy for set size six did differ significantly between the two tasks (Mean accuracy gesture = 0.39, $SD = 0.38$; Mean accuracy dice = 0.08, $SD = 0.18$; $t(20) = -4.02$, $p = 0.001$). Children less commonly said "five and five" for 10, perhaps because both hands showed the same gesture. See Appendix for a full visualization of children's responses for each set size, on each task.

4.3 | Experiment 3 discussion

The results of Experiment 3 suggest that children who are subset-knowers view at least some number gestures holistically prior to learning the cardinal principle. In addition to an overall advantage of labeling number gestures over dot arrays, the greatest difference in accuracy was observed on set size 5. Participants' high accuracy when labeling the 5-gesture (nearly 90% of trials correct on average) is particularly surprising given that all participants in Experiment 3 were subset-knowers and thus could not yet demonstrate an understanding of the meaning of "five" on standard measures of number comprehension.

FIGURE 4 Example stimuli of Consistent Configuration trials (top row) and Inconsistent Configuration trials (bottom row)

5 | EXPERIMENT 4: GESTURE CONFIGURATION VERSUS QUANTITY

Experiments 1–3 show that children's enumeration of number gestures does not follow the same pattern as children's enumeration of shapes. Specifically, when enumerating gestures, accuracy does not dramatically decrease beyond the subitizable range. We attribute this robustness to children's ability to recognize number gestures as a whole, making it unnecessary for them to rely on the more error prone strategy of using the approximate number system to estimate the number of fingers raised. In a final experiment, we created a strong test of this interpretation by presenting children with digitally modified number gestures in which the configuration was inconsistent with the actual number of fingers raised. On key trials (see Figure 4), children saw a canonical five gesture (all fingers and thumb raised) in which one of the fingers had been digitally removed (leaving four fingers), a five configuration but only four fingers; they also saw a canonical four gesture (all fingers raised with thumb down across palm) in which a finger had been digitally added (resulting in five raised fingers), a 4 configuration but with 5 fingers. We reasoned that if children arrive at their answers by using their ANS to estimate the number of individual fingers, they should label gestures in which four fingers are raised as "four" regardless of configuration, and should label gestures in which five fingers are raised as "five" regardless of configuration. Alternatively, if children predominantly use configuration to identify quantity on this task, they should guess "five" when only four fingers are raised as long as the fingers are in a five configuration, and they should guess "four" even when five fingers are raised as long as the fingers are in a four configuration.

5.1 | Methods

5.1.1 | Participants

Thirty-four children (16 female) participated in the study. The mean age was 4.14 years (*SD* = 0.45, range = 2.70–4.82 years). Average family income fell in the $50,000 to $75,000 category. One additional participant was excluded because they were not paying attention to the task and their responses were greater than 10 standard deviations above the mean response of the other participants. However, the results of our main analyses are the same with or without this participant.

5.2 | Procedure

Participants in the current study were obtained from a larger sample of preschool students who participated in a number battery and were sorted into separate studies based on their knower-level. All participants were first asked to count as high as they could and then completed the Give-a-Number task (Wynn, 1990, 1992). Only children who successfully demonstrated knowledge of the cardinal principle were entered into the present study.

5.2.1 | Fast Gestures task (modified)

The Fast Gestures task consisted of a brief familiarization, followed by a single block of test trials. During the familiarization, children were told they were going to see some number gestures and their task was to guess how many fingers they saw raised. They were then shown three examples––a gesture with one finger raised, a gesture with two fingers raised, and a gesture with three fingers raised. The experimenter labeled each gesture with the appropriate number words along with the child. During the test phase, participants were shown a combination of filler trials (canonically configured gestures for 1, 2, and 3), Consistent Configuration trials (canonically configured gestures for 4 and 5) and Inconsistent Configuration trials (4 and 5 gestures in which the configuration of the gesture and the number of fingers raised did not match; see Figure 4). Consistent and Inconsistent Configuration trials were the trials of interest. These seven trials were repeated three times (21 trials total), and within each grouping of seven trials, were presented in a random order.

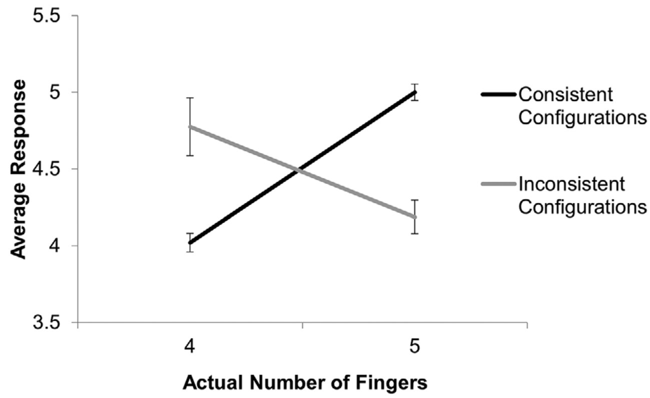

FIGURE 5 Average response by actual number of fingers raised and configuration type. Error bars represent the standard error in participants' estimates for each stimulus

5.3 | Results

First, we looked at the average response children gave for the Consistent Configurations trials and the Inconsistent Configuration trials for set sizes 4 and 5. We ran a 2 (Configuration: Consistent vs. Inconsistent) by 2 (Set Size: 4 vs. 5) ANOVA on participants' average number response. This analysis revealed a main effect of set size ($F(1, 33) = 7.43$, $p = 0.010$, $\eta_p^2 = 0.184$), no effect of configuration ($F(1, 33) = 0.33$, $p = 0.571$, $\eta_p^2 = 0.010$), and a significant set size by configuration interaction ($F(1, 33) = 109.43$, $p < 0.001$, $\eta_p^2 = 0.768$). The average response for each set size and type of configuration is depicted in Figure 5.

As expected, within the consistent configurations, children's estimates for set size 5 were significantly higher than their estimates for set size 4 ($t(33) = 50.00$, $p < 0.001$, $d = 8.57$). In contrast, within the inconsistent configurations, children's estimates for set size 5 were significantly lower than their estimates for set size 4 ($t(33) = -2.13$, $p = 0.041$, $d = 0.69$). Additionally, children's estimates for inconsistently configured set size 5 trials were significantly lower than their estimates for the consistently configured set size 5 trials ($t(33) = -11.52$, $p < 0.001$, $d = 1.98$), and the reverse pattern was found for the estimates of set size 4 ($t(33) = 7.03$, $p < 0.001$, $d = 1.20$).

Across both types of Inconsistent Configuration trials (202 trials total), children's responses matched the configuration on 165 trials (82%), but matched the actual number of raised fingers on only 26 trials (13%), and matched neither the configuration nor the number of fingers on 13 trials (6%). A chi-square goodness-of-fit test confirmed that children's responses were not equally distributed amongst these three categories, X^2 (2, N = 34) = 208.79, $p < 0.001$. Excluding children's responses that matched neither the configuration nor the number of fingers, a binomial test revealed that responses matching the configuration were much more common than responses matching the number of fingers ($p < 0.001$). A second chi-square test found no difference in the distribution of configuration, number, and non-matches between the two Inconsistent Configuration trial types, X^2 (2, $N = 34$) $= 4.41$, $p = 0.110$.

5.4 | Experiment 4 discussion

As a final, strong test of whether children labeled gestures by estimating the number of fingers raised, or by recognizing the handshape of the gesture, Experiment 4 disentangled these two features. The results show that children overwhelmingly answer in accordance with the configuration of number gestures, rather than the actual number of fingers raised. This pattern strongly suggests that, early on, children recognize at least some number gestures as whole gestalts, rather than viewing them as collections of individual fingers

5.5 | Discussion

In a series of experiments, we tested the prediction that children view number gestures purely as nonsymbolic sets of items (e.g., "finger", "finger", "finger") by comparing children's ability to quickly enumerate gestures to their ability to quickly enumerate other nonsymbolic sets (i.e., arrays of shapes). Contrary to this prediction, children's enumeration of gestures in Experiment 1–3 diverged from their enumeration of shapes for quantities beyond the subitizable range (i.e., beyond the 1–3 range), even when shapes were presented in organized arrays (Experiments 2 and 3), and even when children had not yet learned the cardinal principle (Experiment 3). In Experiment 4, we found that children overwhelmingly judged number gestures based on the configuration, rather than the actual number of fingers raised (when the two were engineered to disagree). These finding indicate that, rather than viewing number gestures purely as collections of individual items, children readily view number gestures holistically, akin to how an adult might recognize a numeral or number word.

One intriguing implication of these findings is that number gestures could be considered to be number symbols, even for young children who have not yet learned or only recently learned the cardinal principle. Previous research (e.g., Gunderson et al., 2015) showed that children understand the numerical content of number gestures (i.e., which gesture refers to which quantity) even before they understand the numerical content of corresponding number words. However, our findings go a step farther to show that young children treat number gestures, not as item-based representations, which must be enumerated approximately like dots, but as single-units which can be recognized and labeled with greater accuracy. Thus, connecting the present study to the findings of Gunderson et al. (2015), children may conceive of some number gestures as single units representing specific quantities of items. Participants in our study were not asked to label sets using gestures (as they were in Gunderson et al., 2015). It is therefore possible that children can recognize number gestures as whole gestalts, and understand the numerical content of number gestures, but do not use both pieces of knowledge on the same problems until later in development. Children's previously recorded difficulty with dual representation may make it particularly challenging for them to recognize number gestures' numerical content while also viewing them as symbols (Uttal et al., 2009). Future research is needed to better understand how these various pieces of knowledge develop in relation to each other. Still, the present findings provide strong evidence against the view that young children interpret number gestures like any other nonsymbolic sets of objects and are in line with the possibility that number gestures may be an early form of symbolic number for many children.

Could these results also indicate, as suggested by Nicoladis et al. (2018), that children view number gestures as arbitrary symbols, and initially fail to appreciate their iconicity? In Experiment 4, participants clearly prioritized gestures' form over the actual number of fingers raised. But readers should keep in mind that this was a speeded task and not designed to directly test participants' appreciation of iconicity. Moreover, if the Nicoladis account is true, one still needs to explain why children in Gunderson et al. (2015) labeled sets more accurately with number gestures than with number words if not a result of iconicity. One possible explanation for this gesture advantage that does not involve iconicity is that children are generally exposed to a smaller set of number gestures (one through five and combinations of one through five to make larger sets) than number words, potentially making them easier to learn.

Another interesting question raised by this work is the extent to which number gestures are unique as item-based representations of number that are encoded as whole gestalts. Relatedly, we can ask how number gestures come to be viewed in this way by young children? Previous work has found that slightly older children do treat dice configurations in a way that looks more like how children in our study view number gestures (Jansen et al., 2014). These findings suggest that number gestures are not unique as item-based representations that are viewed holistically. However, children may view number gestures holistically more readily, or at least at an earlier stage of development, than other representations of number (such as dice). Differences in children's experiences with number gestures and dice configurations are likely to contribute to the formation of these differences. Differences in experience could also explain why children were more accurate on some number gestures than others (e.g., five compared to four; though five with its extended thumb is also more distinctive). Beyond differences in experience, there could be other factors that contribute to differences between children's understanding of number gestures and dice; for example, gesture is used more frequently in communication than dice; the fingers in a gesture are all connected via the hand; gestures are embodied representations of number. Future research is needed to examine the extent to which each of these differences contributes to the gesture advantage we observed here.

Despite these remaining questions, our findings contribute to our understanding of how number gestures might facilitate children's acquisition of number words. Early in the process of number development, children connect number gestures to quantities (Gunderson et al., 2015) and to number words (our findings in the present study) more accurately than they connect number words directly to nonsymbolic quantities. This pattern lends support to the possibility that number gestures may serve as a bridge between number words and nonsymbolic representations of quantity (e.g., Gunderson et al., 2015; Orrantia et al., 2022; Wiese, 2003).

Number gestures could also play a supportive role in children's number development by helping children develop numeracy skills that are typically associated with learning number words. For instance, numerical training involving number words and abacus-counting is associated with greater improvements to children's abilities to perform one-to-one correspondence and non-verbal arithmetic, than training involving only non-symbolic numerical comparison (Hyde et al., 2021). To the extent that number gestures also function as symbols, they may have similar beneficial effects on children's numeracy.

In sum, our findings suggest that number gestures share similarities with number words and may even be used as number symbols themselves beginning when children are quite young. This has several important implications. First, it strengthens the basis for comparisons between number gestures and number words, which is useful since comparing children's acquisition of different number systems, such as those in different languages, has been a fruitful pathway for better understanding why learning number words is challenging for young children (e.g., Almoammer et al., 2013; Sarnecka et al., 2007; Wagner et al., 2015). Second, by showing that children more easily (or at least more rapidly) map number words to number gestures than to nonsymbolic sets, our results lend further support to the possibility that number gestures can or do serve as a bridge between nonsymbolic and symbolic number. Finally, our results shed light on the relative difficulty of acquiring different types of knowledge about number gestures and can therefore inform caregivers' and educators' decisions about which skills to focus on during number instruction. Specifically, one may assume that nonverbal number skills (i.e., matching gestures to sets) may precede verbal skills (i.e., matching number words to number gestures), but this assumption is called into question by the ease and accuracy with which children labeled number gestures with number words in our study. Future research should continue to probe the developmental timeline of children's ability to map number gestures to quantities and to number words, while at the same time exploring which of these abilities can usefully be strengthened through interventions to advance children's early number knowledge.

ACKNOWLEDGMENTS

This research was supported by the Hymen Milgrom Supporting Organization to Dominic J. Gibson, Talia Berkowitz, and Susan C. Levine; and by NCER Grant R305B090025 and National Science Foundation Graduate Research Fellowship DGE-1144082 to Dominic Gibson. We thank Jodi Khan for administrative and technical support. This work was completed at the University of Chicago.

CONFLICTS OF INTEREST

We have no conflicts of interest to disclose. This research was approved by The University of Chicago Social and Behavioral Sciences Institutional Review Board and consent was obtained from parents the parents of all participants. The data that support the findings of this study are available from the corresponding author upon reasonable request

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Dominic J. Gibson https://orcid.org/0000-0001-8841-2840

ENDNOTES

[1] Theoretically, we should observe no decrease in accuracy, but even when labeling numerals, children's accuracy decreases as the value of the numeral increases, likely due to differences in familiarity (Jiménez et al., 2017).

[2] Because participants were recruited through a larger investigation into children's number learning and entered into the study depending on their knower-level, we ended up with varying numbers of participants across experiments. However, our main analyses were sufficiently powered, and we observed large effects, which we replicated whenever possible.

REFERENCES

Alibali, M. W., & DiRusso, A. A. (1999). The function of gesture in learning to count: More than keeping track. *Cognitive Development*, *14*(1), 37–56. https://doi.org/10.1016/s0885-2014(99)80017-3

Almoammer, A., Sullivan, J., Donlan, C., Marušič, F., Žaucer, R., O'Donnell, T., & Barner, D. (2013). Grammatical morphology as a source of early number word meanings. *Proceedings of the National Academy of Sciences*, *110*(46), 18448–18453. https://doi.org/10.1073/pnas.1313652110

Crollen, V., Seron, X., & Noël, M. P. (2011). Is finger-counting necessary for the development of arithmetic abilities? *Frontiers in Psychology*, *2*, 242. https://doi.org/10.3389/fpsyg.2011.00242

Di Luca, S., & Pesenti, M. (2011). Finger numeral representations: More than just another symbolic code. *Frontiers in Psychology*, *2*, 272. https://doi.org/10.3389/fpsyg.2011.00272

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615. https://doi.org/10.1016/j.tics.2015.07.013

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., … Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

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

Geary, D. C., vanMarle, K., Chu, F. W., Rouder, J., Hoard, M. K., & Nugent, L. (2018). Early conceptual understanding of cardinality predicts superior school-entry number-system knowledge. *Psychological Science*, *29*(2), 191–205. https://doi.org/10.1177/0956797617729817

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

Gibson, D. J., Gunderson, E. A., Spaepen, E., Levine, S. C., & Goldin-Meadow, S. (2019). Number gestures predict learning of number words. *Developmental Science*, *22*(3), e12791. https://doi.org/10.1111/desc.12791

Goldin-Meadow, S., Levine, S. C., & Jacobs, S. (2014). Gestures role in learning arithmetic. In L. D. Edwards, F. Ferrara & D. Moore-Russo (Eds.), *Emerging Perspectives on Gesture and Embodiment in Mathematics* (pp. 51–72) Information Age Publishing.

Gordon, R., Chernyak, N., & Cordes, S. (2019). Get to the point: Preschoolers' spontaneous gesture use during a cardinality task. *Cognitive Development*, *52*, 100818. https://doi.org/10.1016/j.cogdev.2019.100818

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

Hurst, M., Anderson, U., & Cordes, S. (2017). Mapping among number words, numerals, and nonsymbolic quantities in preschoolers. *Journal of Cognition and Development*, *18*(1), 41–62. https://doi.org/10.1080/15248372.2016.1228653

Huttenlocher, J., Jordan, N. C., & Levine, S. C. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology: General*, *123*(3), 284–296. https://doi.org/10.1037/0096-3445.123.3.284

Hyde, D. C., Mou, Y., Berteletti, I., Spelke, E. S., Dehaene, S., & Piazza, M. (2021). Testing the role of symbols in preschool numeracy: An experimental computer-based intervention study. *PLOS ONE*, *16*(11), e0259775. https://doi.org/10.1371/journal.pone.0259775

Ifrah, G. (2000). The Universal History of Numbers: From Prehistory to the Invention of the Computer. Translated by D. Vellos, E. F. Harding, S. Wood, & I. Monk, New York: John Wiley & Sons.

Jansen, B. R., Hofman, A. D., Straatemeier, M., Bers, B. M., Raijmakers, M. E., & Maas, H. L. (2014). The role of pattern recognition in children's exact enumeration of small numbers. *British Journal of Developmental Psychology*, *32*(2), 178–194. https://doi.org/10.1111/bjdp.12032

Jiménez Lira, C. J., Carver, M., Douglas, H., & LeFevre, J. A. (2017). The integration of symbolic and non-symbolic representations of exact quantity in preschool children. *Cognition*, *166*, 382–397. https://doi.org/10.1016/j.cognition.2017.05.033

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *The American Journal of Psychology*, *62*(4), 498–525. https://doi.org/10.2307/1418556

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*(2), 395–438. https://doi.org/10.1016/j.cognition.2006.10.005

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*(1), 1–22. https://doi.org/10.1037/0096-3445.111.1.1

Mix, K. S. (2008). Surface similarity and label knowledge impact early numerical comparisons. *British Journal of Developmental Psychology*, *26*, 13–32. https://doi.org/10.1348/026151007X189109

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Namy. (2008). Recognition of iconicity doesn't come for free. *Developmental Science*, *11*(6), 841–846. https://doi.org/10.1111/j.1467-7687.2008.00732.x

Nicoladis, E., Marentette, P., Pika, S., & Barbosa, P. G. (2018). Young children show little sensitivity to the iconicity in number gestures. *Language Learning and Development*, *14*(4), 297–319. https://doi.org/10.1080/15475441.2018.1444486

Orrantia, J., Muñez, D., Sanchez, R., & Matilla, L. (2022). Supporting the understanding of cardinal number knowledge in preschoolers: Evidence from instructional practices based on finger patterns. *Early Childhood Research Quarterly*, *61*, 81–89. https://doi.org/10.1016/j.ecresq.2022.05.009

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: From a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130300. https://doi.org/10.1098/rstb.2013.0300

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

Sarnecka, B. W., Kamenskaya, V. G., Yamana, Y., Ogura, T., & Yudovina, Y. B. (2007). From grammatical number to exact numbers: Early meanings of 'one', 'two', and 'three' in English, Russian, and Japanese. *Cognitive Psychology*, *55*(2), 136–168. https://doi.org/10.1016/j.cogpsych.2006.09.001

Sarnecka, B. W., & Lee, M. D. (2009). Levels of number knowledge during early childhood. *Journal of Experimental Child Psychology*, *103*(3), 325–337. https://doi.org/10.1016/j.jecp.2009.02.007

Uttal, D. H., O'Doherty, K., Newland, R., Hand, L. L., & DeLoache, J. (2009). Dual representation and the linking of concrete and symbolic representations. *Child Development Perspectives*, *3*(3), 156–159. https://doi.org/10.1111/j.1750-8606.2009.00097.x

Wagner, K., Kimura, K., Cheung, P., & Barner, D. (2015). Why is number word learning hard? Evidence from bilingual learners. *Cognitive Psychology*, *83*, 1–21. https://doi.org/10.1016/j.cogpsych.2015.08.006

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science*, *10*(2), 130–137. https://doi.org/10.1111/1467-9280.00120

Wiese, H. (2007). The co-evolution of number concepts and counting words. *Lingua*, *117*(5), 758–772.

Wiese, H. (2003). Iconic and non-iconic stages in number development: The role of language. *Trends in Cognitive Sciences*, *7*(9), 385–390. https://doi.org/10.1016/S1364-6613(03)00192-X

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

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

How to cite this article: Gibson, D. J., Berkowitz, T., Butts, J., Goldin-Meadow, S., & Levine, S. C. (2022). Young children interpret number gestures differently than nonsymbolic sets. *Developmental Science*, e13335. https://doi.org/10.1111/desc.13335

APPENDIX

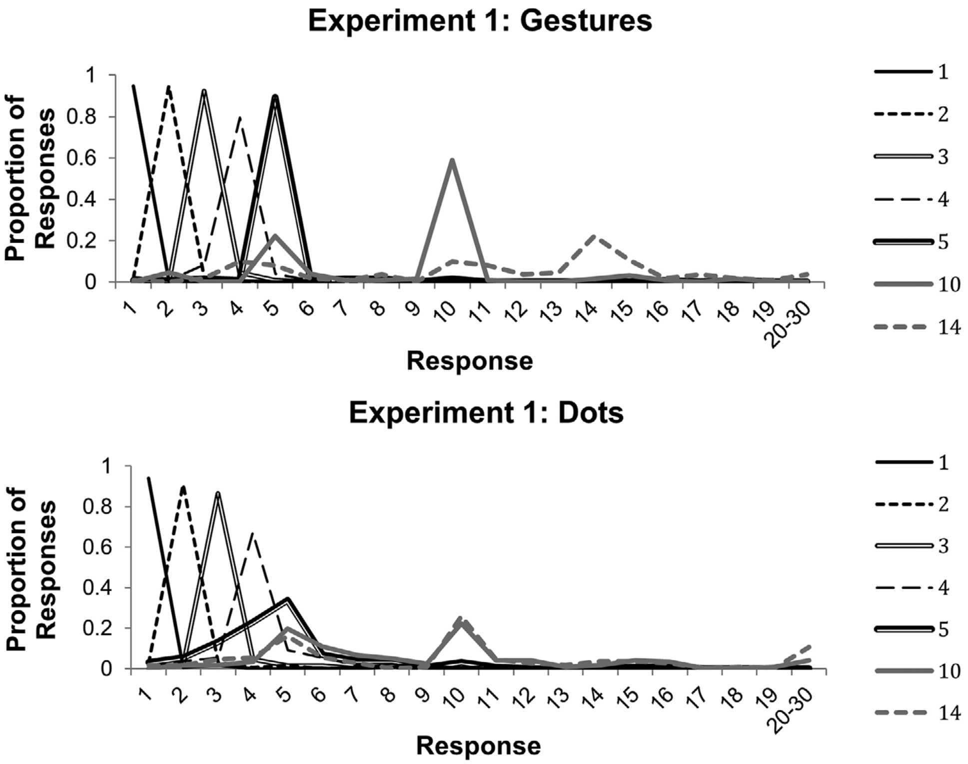

FIGURE A1 Distribution of responses for each set size on the Fast Gestures task (top) and Fast Dots task (bottom). Each line represents a target set size and shows the distribution of actual responses children made for that set size on each task. For the Fast Dots task, the lower peaks and wider distribution of responses as set sizes increase reflects children's reliance on the Approximate Number System for higher set sizes (e.g., 5). In contrast, the distribution of responses for set sizes 4 and 5 in the Fast Gestures task look remarkably similar to those for set sizes 1–3.

Developmental Science

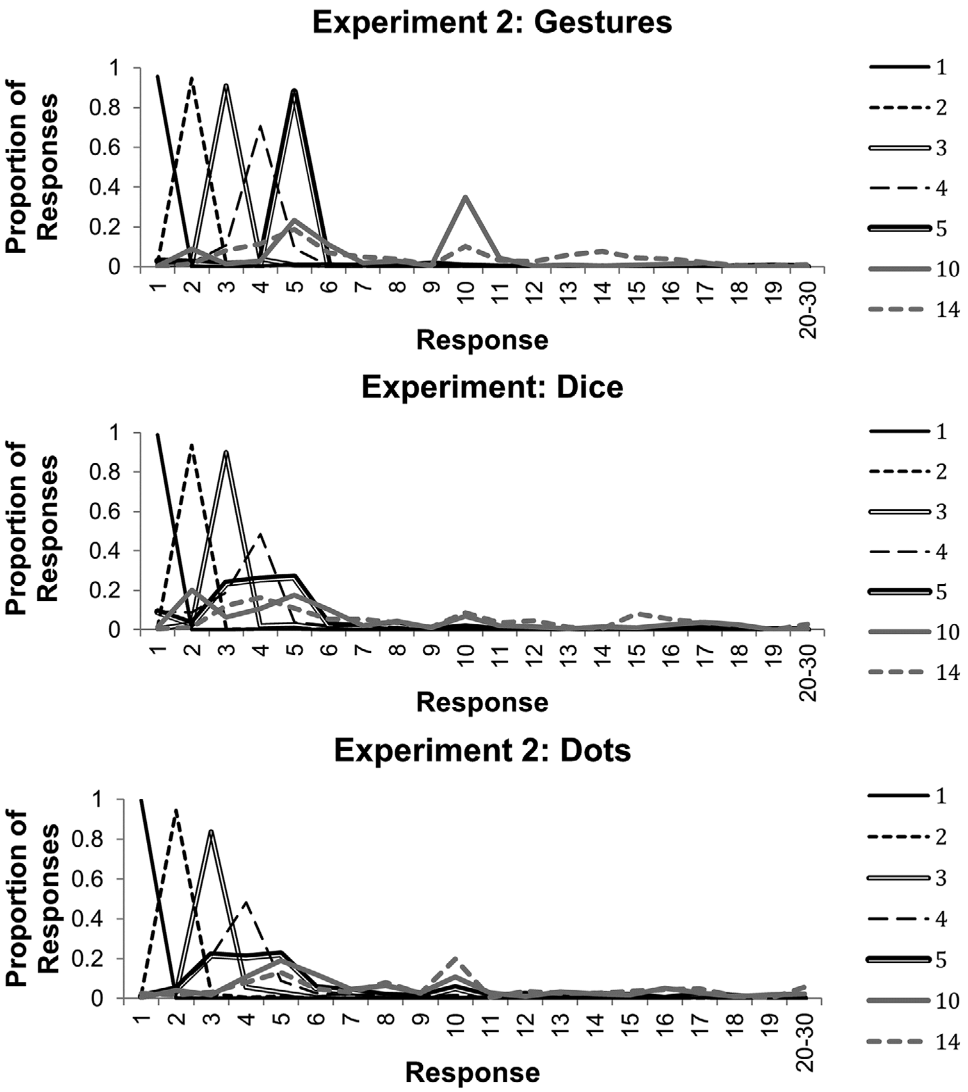

FIGURE A2 Distribution of responses for each set size on the Fast Gestures task (top), Fast Dice task (middle) and Fast Dots task (bottom). Each line represents a target set size and shows the distribution of actual responses children made for that set size on each task. Like the results of the Fast Dots task in Experiment 1, the distributions of children's responses on the Fast Dots and Fast Dice tasks get flatter as set sizes increase. In contrast, as in Experiment 1, the distribution of responses for set sizes 4 and 5 in the Fast Gestures task look remarkably similar to those for set sizes 1–3.

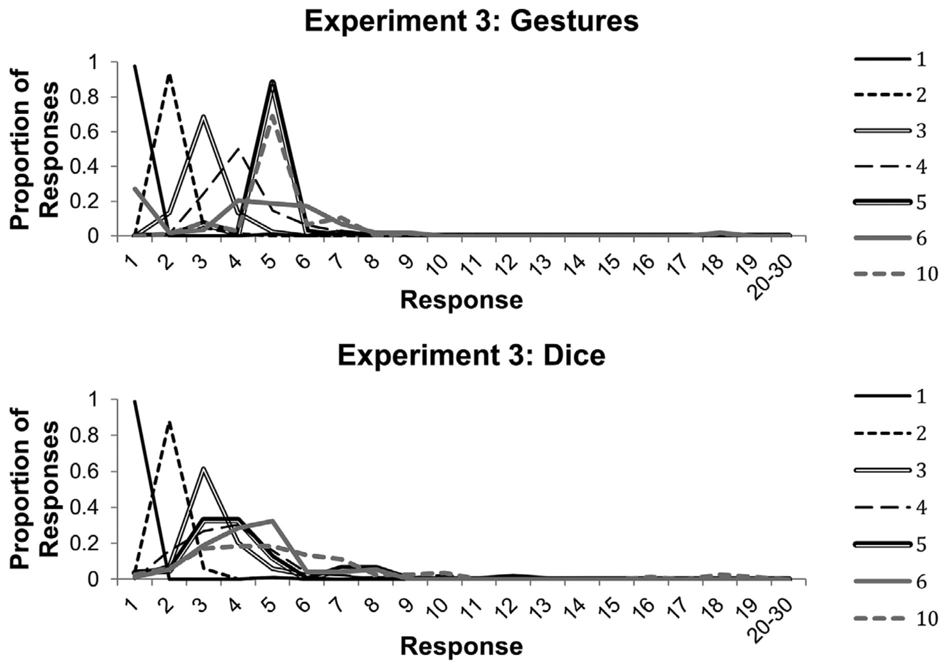

FIGURE A3 Distribution of responses for each set size on the Fast Gestures task (top) and Fast Dice task (bottom) and Fast Dots task (bottom). Each line represents a target set size and shows the distribution of actual responses children made for that set size on each task. In addition to the overall greater accuracy of children's responses on the Fast Gestures task, particularly for set size 5, children's errors look quite different in the two tasks. For instance, on set size 10 for the Fast Gestures task, children frequently responded "five".

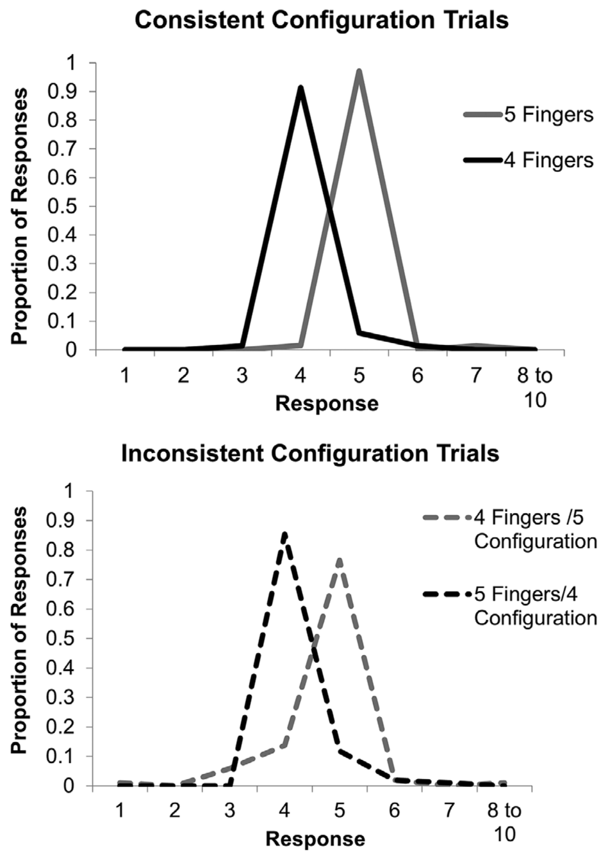

FIGURE A4 Distribution of responses for each set size (represented by individual lines) on Consistent Configuration trials (top) and Inconsistent Configuration trials (bottom).