CHILD DEVELOPMENT

Child Development, May/June 2021, Volume 92, Number 3, Pages 1011–1027

First and Second Graders Successfully Reason About Ratios With Both Dot Arrays and Arabic Numerals

Emily Szkudlarek and Elizabeth M. Brannon
University of Pennsylvania

Children struggle with exact, symbolic ratio reasoning, but prior research demonstrates children show surprising intuition when making approximate, nonsymbolic ratio judgments. In the current experiment, eighty-five 6- to 8-year-old children made approximate ratio judgments with dot arrays and numerals. Children were adept at approximate ratio reasoning in both formats and improved with age. Children who engaged in the nonsymbolic task first performed better on the symbolic task compared to children tested in the reverse order, suggesting that nonsymbolic ratio reasoning may function as a scaffold for symbolic ratio reasoning. Nonsymbolic ratio reasoning mediated the relation between children's numerosity comparison performance and symbolic mathematics performance in the domain of probabilities, but numerosity comparison performance explained significant unique variance in general numeration skills.

Ratio reasoning is prevalent in everyday life. We reason about ratios when we estimate how long it will take to walk to the grocery store, decide how many hands we need to carry all the grocery bags into the house, or when we slice a birthday cake to ensure that everyone at the party will get a piece. Mathematics provides a symbolic system to calculate exact solutions to these ratio reasoning problems, however, in everyday life we typically avoid such precise calculations and instead arrive at an approximate, workable solution. These approximate ratio reasoning skills are present from infancy and are shared with other animal species. Comparative and developmental research demonstrates that animals and human babies possess an intuitive capacity for ratio reasoning that is not dependent on language or knowledge of symbolic mathematics. For example, when an ape is given a choice between two human hands that each randomly drew one food item from one of two buckets, the ape will tend to pick the hand that drew from the bucket that contained the more favorable ratio of banana pellets to carrots (Eckert, Call, Hermes, Herrmann, & Rakoczy, 2018; Rakoczy et al., 2014). Similarly, rhesus macaques can be trained to pick the array with the more favorable ratio of a rewarded shape, or to match different colored line lengths based on ratio, for a juice reward (Drucker, Rossa, & Brannon, 2015; Vallentin & Nieder, 2008). When 6-month-old human infants are habituated to a particular ratio of yellow to blue objects within an array, they look longer at a novel compared to a familiar ratio (McCrink & Wynn, 2007). Ten to 12-month-old infants will reliably crawl toward a cup that contains a lollipop drawn from a bucket with a more favorable ratio of their preferred color lollipop (Denison & Xu, 2014). Relatedly, it violates infants' expectations when a sample is drawn from a population of items that does not reflect the population distribution (Kayhan, Gredebäck, & Lindskog, 2017; Xu & Garcia, 2008). For example, when 8-month-old infants see a box filled with ping-pong balls where the majority is red and only a few are white, they look longer when four of the five balls randomly drawn from this box are white (Xu & Garcia, 2008). Thus, infants and nonhuman primates are equipped with intuitive ratio reasoning skills that are flexible to probabilistic context and the

We thank Luis Rosario, Mikie Sakanaka, Joseph Dembik, Lauren Paulson, I'Mani Sellers and Yvette Almaguer for help with data collection, and Nicholas K DeWind and Stephanie Bugden for helpful discussions about the manuscript. We also thank the parents and after school programs that participated in this research. The research reported here was supported by an R01 HD079106 award to Elizabeth M. Brannon, a National Science Foundation Division of Research on Learning Award #1660973 to Elizabeth M. Brannon, and a Ruth L. Kirschstein Predoctoral Individual National Research Service Award F31 HD095579-01 to Emily Szkudlarek.

Correspondence concerning this article should be addressed to Emily Szkudlarek, Wisconsin Center for Education Research, University of Wisconsin-Madison, 1025 West Johnson St., Room 698, Madison, WI 53706. Electronic mail may be sent to szkudlarek@wisc.edu.

 0009-3920/2021/9203-0016
DOI: 10.1111/cdev.13470

depiction of the ratio. An important question is whether this powerful ratio reasoning mechanism can be harnessed to improve symbolic ratio reasoning ability in school-age children.

Children continue to reason nonsymbolically and approximately with ratios before they are capable of calculating with symbolic ratio representations. Six-year-old children can solve proportional reasoning problems displayed with continuous magnitudes (e.g., line lengths), and by age 10 they can solve the same ratio problems presented with discrete magnitudes (e.g., squares; Boyer & Levine, 2015; Boyer, Levine, & Huttenlocher, 2008). Elementary and preschool-aged children (6–11 years old) can successfully pick which of two jars has the more favorable ratio of their preferred color object (Falk, Yudilevich-Assouline, & Elstein, 2012; Yost, Siegel, & Andrews, 1962). Moreover, there is a connection between a student's nonsymbolic proportional reasoning ability and their symbolic math skill. A student's nonsymbolic proportional reasoning ability in fifth grade predicts unique variance in knowledge of symbolic fractions in sixth grade (Jordan, Resnick, Rodrigues, Hansen, & Dyson, 2016). Performance on a spatial proportional reasoning task is correlated with fraction understanding in 8- to 10-year-olds (Möhring, Newcombe, Levine, & Frick, 2016). In sum, children possess an intuitive grasp of nonsymbolic ratio magnitudes that is linked to their symbolic math skill. But in contrast to children's prodigious intuitive ratio reasoning skill, exact, symbolic ratio reasoning is very difficult for children to acquire.

According to the US Common Core standards, children begin formal education about ratio in third grade, but by fourth grade only 32% of students is able to identify whether simple fractions are greater or < ½, and only 27% of eighth graders can successfully place a rational number in the correct spot on a number line (National Center for Education Statistics, 2017). Ratios, fractions, and proportions are critically important for advanced math education, thus deficits in symbolic ratio reasoning are a major challenge for math educators (Butterworth, Varma, & Laurillard, 2011; Duncan et al., 2007; Parsons & Bynner, 2005; Siegler et al., 2012). One possible way to improve children's symbolic ratio reasoning is to use a child's intuitive understanding of ratio to scaffold their symbolic learning (Ahl, Moore, & Dixon, 1992; Falk & Wilkening, 1998; Fujimura, 2001). Prior research has established that children are able to nonsymbolically and approximately calculate over nonsymbolic dot arrays by adding, subtracting, multiplying, dividing, solving for unknown values in algebra problems, and placing dot arrays on a number line (Barth et al., 2006; Honoré & Noël, 2016; Kibbe & Feigenson, 2017; McCrink, Shafto, & Barth, 2016; Park, Bermudez, Roberts, & Brannon, 2016). Moreover, training with some nonsymbolic and approximate math operations yields improvement in symbolic math ability. For example, practice of adding and subtracting arrays of objects improves the general math skill of preschoolers after 10 days of training (Park et al., 2016; Szkudlarek & Brannon, 2018). Similarly, practice with a symbolic and nonsymbolic number line task improves number line and arithmetic performance in 8- to 10-year-old children (Kucian et al., 2011). It is therefore possible that practice with nonsymbolic ratio reasoning could benefit symbolic ratio reasoning in a similar way.

The first goal of this study was to examine the nonsymbolic, approximate ratio comparison skills of children in early elementary school, and to test whether children can extend their nonsymbolic, approximate ratio reasoning ability to a novel *symbolic*, approximate ratio comparison task. To this end, we created both a nonsymbolic and symbolic version of the same approximate ratio reasoning task. On each trial, children were presented with an illustration of two gumball machines both filled with blue and white colored gumballs (dots) or blue and white numerals (Figure 1). The task was to pick the machine with the best chance of producing a blue (or white, counterbalanced) gumball. Crucially, the task did not require exact calculation of the ratio of ratios, but instead only required children to identify which ratio of blue to white was more favorable. Due to the approximate nature of the task, we predicted that children would perform above chance despite their lack of formal knowledge of ratio or probabilistic reasoning with both formats of the task (e.g., Gilmore, McCarthy, & Spelke, 2007 for a similar finding with approximate addition and subtraction). We further hypothesized that if children were able to connect their nonsymbolic ratio reasoning skills to the symbolic version of the calculation, engaging in nonsymbolic ratio reasoning should provide a foothold for children as they attempt to make the isomorphic judgment with symbols. To test this, each child completed both the symbolic and nonsymbolic ratio comparison tasks with order counterbalanced across children. If children's symbolic ratio reasoning can be scaffolded by engaging in the nonsymbolic ratio reasoning task, then performance on the symbolic task should be modulated by the order of the two tasks.

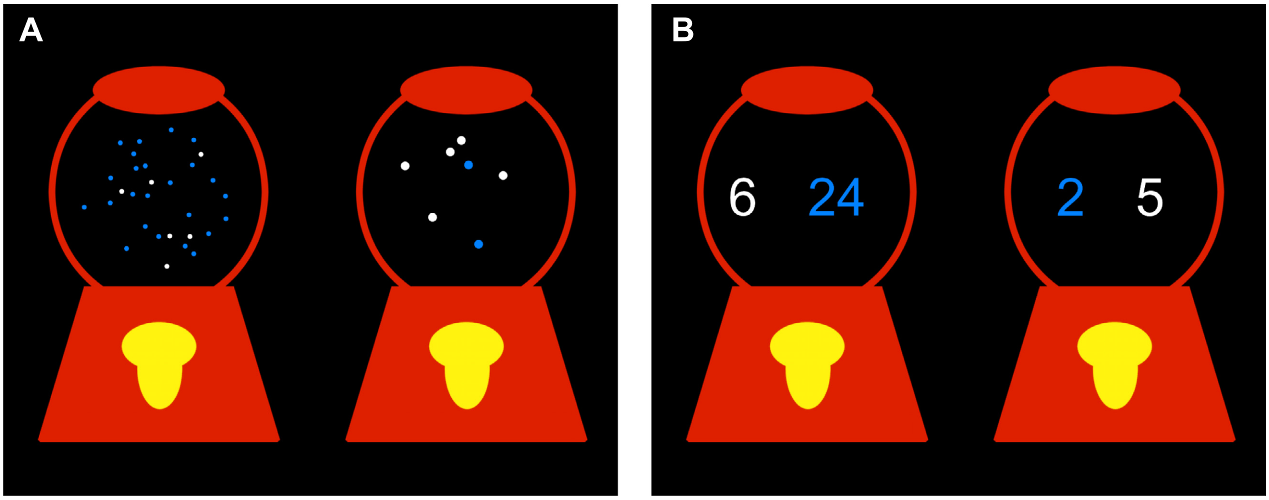

Figure 1. Illustration of one trial of the (A) nonsymbolic ratio comparison task and (B) symbolic ratio comparison task. [Color figure can be viewed at wileyonlinelibrary.com]

The second goal of this study was to examine the degree to which children made true ratio comparisons or instead relied on simpler unidimensional heuristics. Prior work found that when children 7 years of age or younger were presented with two ratios they tend to erroneously compare only the number of preferred items in an integer-based strategy (Clarke & Roche, 2009; Falk et al., 2012; Jeong, Levine, & Huttenlocher, 2007; Obersteiner, Bernhard, & Reiss, 2015; Shaklee & Paszek, 1985; Siegler, Strauss, & Levin, 1981). As children progress in their mathematical knowledge, they are more likely to perform a true ratio comparison and consider both the preferred and nonpreferred items in each array. However, previous research on ratio comparison strategies has relied on self-report or differences in accuracy between trial types to test for the use of a particular heuristic. These approaches have serious drawbacks. First, self-report may be inaccurate, especially among young children who have not yet learned formal ratio reasoning in school. Second, a comparison of accuracy by trial type does not account for the fact that many binary choice ratio comparisons can be solved using more than one strategy. For example, imagine two gumball machines where the left machine has 25 blue gumballs and 15 white, and the right machine has 12 blue gumballs and 18 white. A child needs to pick the machine with the best chance of getting a blue gumball on their first try. The correct answer based on a comparison of ratios is the left machine. But a child could also pick the left machine because it has a larger total number of gumballs, because it has the largest absolute number of blue gumballs, or the smallest absolute quantity of white gumballs. Greater accuracy on trials of this type does not disambiguate between the use of these strategies. A more flexible analysis of strategy is needed to describe the potential use of multiple strategies within one subject. Here, we constructed the stimuli such that the correct answer was orthogonal to the total number of items in a gumball machine, and the correlation between the correct choice and the one with more of the preferred or nonpreferred colored items was minimized. We then modeled the left or right choices a child would make under three alternative unidimensional heuristics that children could use on a binary choice ratio comparison task. This stimulus structure allowed us to detect the degree to which children used unidimensional heuristics or a ratio comparison to solve the ratio tasks by looking at a child's pattern of behavior across all trials. Our strategy analysis had three goals. The first was to determine whether children can perform a true ratio comparison independent of incorrect heuristics on both the symbolic and nonsymbolic tasks. The second was to examine whether the errors children make during the ratio tasks correspond with using an incorrect heuristic. The third was to assess the degree to which task order (nonsymbolic or symbolic first) affects not only overall accuracy, but the strategies children employ to complete the task. We hypothesized that engaging in nonsymbolic ratio comparison would shift the strategies children use to solve the symbolic ratio task and lead to improved symbolic ratio reasoning. More specifically, we expected that children might rely less on heuristics in the symbolic version of the task if they first engaged in the nonsymbolic task.

The third goal of this study was to determine whether the Approximate Number System (ANS) is a cognitive foundation underlying nonsymbolic ratio reasoning. The ANS supports the ability to represent numerical magnitude without symbols, and is present in infants, adults, and many animal species (Feigenson, Dehaene, & Spelke, 2004). Previous work suggests that the ANS is involved in binary choice ratio comparison tasks, because performance on these tasks is dependent on the ratio of ratios being compared, as predicted by *Weber's* law (Drucker et al., 2015; Eckert et al., 2018; McCrink & Wynn, 2007). Ratio-dependent numerical discrimination following Weber's law is a hallmark of the ANS (Feigenson et al., 2004). Moving beyond the description of a ratio effect, in this study, we independently measured each child's ANS acuity as a stronger test of whether children with sharper ANS acuity are better at nonsymbolic ratio reasoning.

We also examined the relation between ratio reasoning skill, ANS acuity, and formal math ability. A multitude of studies has linked individual differences in the acuity of ANS representations to a variety of symbolic math skills, but the mechanism for this relation is unknown (Halberda, Mazzocco, & Feigenson, 2008; see Schneider et al., 2016 for meta-analysis). Recent work suggests that accuracy in performing mathematical operations using nonsymbolic quantities may be a better predictor of symbolic math ability than ANS acuity (Matthews, Lewis, & Hubbard, 2016; Pinheiro-Chagas et al., 2014; Starr, Roberts, & Brannon, 2016). Indeed, among university undergraduates, nonsymbolic ratio comparison accuracy was a stronger predictor than ANS acuity of a variety of symbolic math tasks. This finding led to the proposal of a Ratio Processing System (Matthews et al., 2016). It is possible that the Ratio Processing System, when dealing in discrete magnitudes, functions as a "higher order extension" of the ANS (Lewis, Matthews, & Hubbard, 2016). Here we test the possibility that performing a nonsymbolic operation, in this case specifically a ratio comparison operation, is a mechanism of the relation between ANS acuity and symbolic math. We hypothesize that sharper ANS acuity allows for better nonsymbolic ratio calculation. In turn, better nonsymbolic ratio calculation allows children to conceptually ground symbolic ratio representations in their nonsymbolic sense of ratio, leading to better symbolic calculation skill. To test this, we asked whether nonsymbolic ratio skill mediates the relation between ANS acuity and symbolic math skill as measured by our symbolic ratio comparison test and two subtests of the Key-Math-3 standardized test. The Key-Math-3 test is divided into subtests that represent different symbolic math skills (numeration, algebra, geometry, measurement, data analysis and probability). Based on our hypothesis that nonsymbolic ratio calculation leads to a conceptual grounding of ratio computation, we administered two subtests with the goal of targeting both ratio based and nonratio-based symbolic math concepts. The Data Analysis and Probability section includes questions about probability and graphical representations of data; these math concepts overlap with the concepts involved in a nonsymbolic probabilistic ratio comparison judgment. We also administered the Numeration section, which is a test of general counting and calculation skill. These general math skills have less conceptual overlap with nonsymbolic ratio comparison. Thus, we hypothesize that nonsymbolic ratio skill will mediate the relation between ANS acuity and questions about Data Analysis and Probability but will not mediate the relation between ANS acuity and general Numeration skill.

To summarize, the present experiment was designed with three major goals. First, to determine whether children's nonsymbolic ratio reasoning skill functions to scaffold symbolic ratio reasoning. Second, to examine the degree to which children utilize true ratio reasoning or instead rely on incorrect unidimensional heuristics to solve a binary choice ratio comparison task. Third, to investigate the relations between ANS acuity, nonsymbolic ratio reasoning, and symbolic math performance. Taken together, these questions have important implications for the value of including nonsymbolic ratio calculation in early math education.

Method

Subjects

Eighty-five 6- to 8-year-old children were tested (22 six-year-olds, 33 seven-year-olds, 30 eight-year-olds, $M_{\text{age}} = 7.6$ years old, $SD = 0.83$ years; 35 female, 47 male, 3 chose not to report, 5 kindergarteners, 35 first graders, 31 second graders, 12 third graders, 2 unknown). We chose this age range because we wanted to test children's knowledge of ratio before they begin to learn formally about fraction concepts in school. Written parental consent and children's verbal assent were obtained in accordance with a protocol accepted by the University of

Pennsylvania's Institutional Review Board. Ten additional children were consented but were excluded from the final sample because they did not complete both the nonsymbolic and symbolic ratio comparison tasks. The parents of 43 children in our sample elected to fill out a detailed demographics questionnaire. Of this subset of the sample, 37% identified as Hispanic or Latino, 49% identified as African American, 21% as Caucasian, 5% as Asian, 12% as more than one race, and 13% chose not to report. Our sample included a large proportion of children from families with household incomes of 50,000 or less (23% \$0–\$25,0000, 42% \$25,000–\$50,000, 5% \$50,000–\$75,000, 2% \$75,000–\$100,000, 7% \$100,000–\$150,000, 12% \$150,000+, and 9% chose not to report). All subjects were recruited from six after school programs in the Philadelphia, PA area between October 2016 and April 2017. A subset of the children who completed both the nonsymbolic and symbolic ratio comparison tasks completed additional assessments (Dot comparison, $n = 75$; Key-Math Numeration assessment, $n = 78$; the Key-Math Data Analysis and Probability assessment, $n = 73$; the Woodcock–Johnson Basic Reading Skills cluster, $n = 72$; and a measure of numeral identification, $n = 82$). All participants received a small toy as a thank you gift after completion of the experiment.

Procedure

Children completed all tasks individually with an experimenter in a quiet room at their after-school program. Children completed the nonsymbolic and symbolic ratio comparison tasks first, and the order of the tasks was counterbalanced across children (nonsymbolic first $n = 40$, symbolic first $n = 45$). All but nine participants received the two ratio comparison tasks on the same day. The order in which all other tasks were administered was random across participants and was dependent on the child's availability. Each participant was tested for a total of 45–60 min across 2–3 days. Children also completed a math anxiety questionnaire, but these results are not included in this paper. Children received stickers throughout the session to maintain motivation.

Experimental Tasks

Both the symbolic and nonsymbolic tasks were run in MATLAB and programmed using the Psychophysics Toolbox extension (Brainard, 1997; Kleiner, Brainard & Pelli, 2007; Pelli, 1997). The programs were run on a 15-in. touch screen laptop computer. We have posted the ratio comparison task scripts and data at this link https://osf.io/2fkbn/.

Introduction to the Ratio Comparison Tasks

All children regardless of whether they were given the symbolic or nonsymbolic version of the ratio task first were introduced to an alien character next to a single red and yellow gumball machine that held 11 orange colored gumballs (dots). Children were shown that only one gumball comes out of the machine at a time and that the alien doesn't care if he gets a big or small gumball. The child was also told that all of the gumballs regardless of their placement in the machine have an equal chance of coming out.

Nonsymbolic Ratio Comparison Task

Children were given five nonsymbolic practice trials. On the first trial, children were presented with two gumball machines where one machine contained 10 green gumballs and the other contained 10 orange gumballs. The child was told that the alien's favorite color gumball is orange. As the experimenter pointed to each machine the child was told "There are this many green gumballs in this machine and this many orange gumballs in this machine." The child was then asked "Which machine gives Mr. Alien the best chance of getting an orange gumball on his first try?" The child was asked to touch the gumball machine of their choice. Children were given verbal feedback by the experimenter. The remaining four practice trials all had mixed green and orange gumballs in each machine. In one trial the correct choice had more orange gumballs (8:2 vs. 2:8), in one trial the correct choice had more orange gumballs and a greater number of total gumballs (6:4 vs. 1:4), in one trial the correct answer had fewer total items (2:3 vs. 2:8), and in the last trial the correct answer had fewer orange gumballs (3:2 vs. 4:6). Regardless of accuracy on this last trial, the child was told "See, the machine with the *most* orange gumballs is not always the one with the *best chance* of an orange gumball." On all practice trials, the experimenter did not proceed until the child picked the correct machine. After the practice trials, children were told that the alien wants a blue or white gumball, counterbalanced across participants. Children completed 60 trials. After a correct response, children saw three smiling aliens, a spaceship blasting off, and the word

"Great!". After an incorrect response, children saw a spaceship crashing into a planet, and the words "Let's try again!". The only instructions a child was given throughout testing were "Which machine has the best chance of giving a (blue/white) color gumball?". Children were instructed not to count the dots if they appeared to be doing so, and to respond as quickly as possible. The dots moved altogether in a circular motion on the screen for 300 ms at the beginning of each trial to encourage interest. There were 30 unique ratio comparisons, and each ratio comparison occurred twice. The number of blue or white gumballs within one machine ranged from 1 to 30. The ratios of ratios (ratio of preferred to nonpreferred in machine 1/ratio of preferred to nonpreferred in machine 2) ranged from 1.5 to 10. The correct answer was orthogonal to the total number of items in a gumball machine (no correlation), and there was a small correlation between the correct choice and the one with the greater number of the preferred color (r = .2) and between the correct choice and the one with less of the nonpreferred items (r = .2). The stimuli are provided in Appendix S1 (Table A2). The stimuli were also counterbalanced to control for congruency with additive reasoning (Obersteiner et al., 2015). Specifically, at each ratio of ratio level there were an equal number of trials where both machines, only one machine, or neither machine had a greater number of the preferred compared to the nonpreferred color (Figure A2 in Appendix S1).

Symbolic Ratio Comparison Task

The procedure and numerical values were identical to those described for the nonsymbolic task. In the symbolic version of the task, the number of gumballs in each gumball machine was represented by blue and white Arabic numerals 1 to 30, instead of dots (see Figure 1). The introduction to Mr. Alien and how the gumball machine works were shown with gumballs (dots) in the machine to exactly match the nonsymbolic version of the task. The five practice trials used green and orange numerals, and as in the nonsymbolic version, the experimenter did not use number words when pointing to the contents of the gumball machines.

Dot Comparison Task

Two dots arrays appeared on a black screen for 750 ms. The dot arrays were subsequently occluded and the child was required to touch the array with a greater quantity. Children completed 200 trials of this task, with feedback on every trial. The number of dots ranged from 8 to 32. The stimuli were created to evenly sample a stimulus space that varied by the ratio between the number, size, and the spacing of the dots. To encourage greater reliability of the measurement, trial level difficulty was titrated (Lindskog, Winman, Juslin, & Poom, 2013). The titration procedure calculated the percentage correct over the last five trials. The ratio between the two dot arrays moved one log level farther apart if the accuracy was less than 70% and moved one log level closer together if the accuracy was > 80%. A quantitative index of each child's ANS acuity was calculated with the Weber fraction (w) as specified in (DeWind, Adams, Platt, & Brannon, 2015). This model accounts for the effects of nonnumerical features of dot arrays on numerical discrimination (DeWind et al., 2015).

Numeral Identification Task

Each child's numeral recognition ability was assessed by presenting the numerals 1–30 individually on index cards. The numerals were displayed in random order, and the child was asked "What number is this?" The accuracy of each child's response was recorded.

Key Math-3 Diagnostic Assessment

The Numeration and Data Analysis and Probability sections of the Key Math-3 Diagnostic Assessment Form B (Connolly, 2007) were administered. The Numeration section is a test of general basic math skills such as place value, counting, the relative magnitude of numbers, and an understanding of fractions, decimals, and percentages. For example, children are presented with four numerals and told "Read the numbers in order, from least to greatest". While the Numeration subtest does contain questions about fractions, the vast majority of these questions occurred farther in the test than was age-appropriate for the children in our sample to reach. The Data Analysis and Probability section targets math content related to concepts of probability, statistics, and graphical representations of data. For example, this section includes questions where children see relevant pictures and are asked "Which spinner gives you an equal chance of landing on green or white?" and "Here is a picture graph of the animals in a pet store. There are two more turtles than which animal?" While some of these questions include nonsymbolic elements (animal pictures, spinners, tally marks) they require an exact answer in response to formalized, symbolic math language. In this way we consider the Data & Analysis and Probability section to be an age-appropriate symbolic math measure that tests formal math concepts related to our nonsymbolic ratio test. We used the age-standardized scale score for both sections of the test.

Woodcock–Johnson IV Test of Cognitive Abilities

Participants' reading abilities were assessed using the "Basic Reading Skills" cluster of the Woodcock–Johnson. This cluster is comprised of the "Letter-Word Identification" and "Word Attack" subtests. In the "Letter-Word Identification" subtest, participants named letters and read words aloud. In "Word Attack," participants read nonsense words and identified letter sounds. We used the age-standardized Basic Reading Skills score.

Analysis Plan

Due to the novel nature of our symbolic, approximate ratio comparison task and our hypothesis on the relation between ANS acuity and nonsymbolic ratio calculation all analyses presented are exploratory.

Strategy Analysis

We examined three incorrect heuristics: picking the machine with a greater absolute number of the preferred color gumballs, picking the machine with a greater number of total items, and picking the machine with less of the nonpreferred color gumball. We will refer to these strategies as the "More Good", "More Items", and "Less Bad" strategies, respectively. Three subjects were excluded from these analyses because they did not complete the full 60 trials for one of the two ratio tasks.

We first identified for each trial whether the correct ratio strategy would result in a left machine response (coded as 0) or a right machine response (coded as 1). We refer to this as the Ratio Model. For example, imagine a trial with three white and seven blue gumballs in the left machine and two white and 14 blue gumballs in the right machine where white is the preferred color. The Ratio Model would indicate a 0 since the left machine has a higher ratio of the preferred color.

To test whether children performed true ratio comparisons independent of incorrect heuristics we constructed three models for each of the three heuristics. The Heuristic Model indicated the left and right responses aligned with each heuristic (Model 1). The Child's Deviation from Heuristic Model indicated how each child's actual choices deviated from each Heuristic Model (Model 2). The Ratio Deviation from Heuristic Model indicated how the Ratio Model and each Heuristic Model differ (Model 3).

To create each Heuristic Model we identified the left or right response a child would make using each heuristic (1 for the choice of the right machine, and 0 for a choice of the left machine). For the example trial described earlier, the More Good and Less Bad Heuristic Models would both indicate the choice of the left machine (coded with a 0) for that trial, whereas the More Items Heuristic Model would indicate the choice of the right machine (coded with a 1).

To create each Child's Deviation from Heuristic Model we subtracted each Heuristic Model from children's actual left and right responses on the task and took the absolute value. A 0 indicated that a child's response corresponded to the incorrect heuristic, and a 1 indicated their response differed from that predicted by the given Heuristic Model.

To create each Ratio Deviation from Heuristic Model we subtracted each Heuristic Model from the Ratio Model and took the absolute value. Thus, for each heuristic, the Ratio Deviation from Heuristic Model indicated whether the heuristic would predict the same response as the Ratio Model (coded as 0) or a different response (coded as 1). To continue with our example trial (3 white and 7 blue on left, 2 white and 14 blue on right), the Ratio Deviation from Heuristic Model for the More Good and Less Bad heuristics indicates 0 for this trial because a child using either heuristic would pick the correct machine. However, the Ratio Deviation from Heuristic Model for the More Items heuristic indicates 1 because the More Items heuristic predicts the opposite response from the Ratio Model for this trial.

We next calculated a Pearson correlation coefficient between the Child's Deviation from Heuristic Model and the Ratio Deviation from Heuristic Model for each child. A perfect correlation between these models ($r = 1$) would indicate perfect ratio comparison performance or 100% accuracy. If a child was exclusively using one of the incorrect heuristics, then the Child's Deviation from Heuristic Model would indicate all 0s (no deviations from using a heuristic) and so the correlation between the two models would be undefined because the Child's Deviation from Heuristic Model would have a standard deviation of zero. Perfect heuristic use with random noise will thus generate correlation coefficients close to zero. This analysis results in six correlation coefficients for each subject, one for each heuristic on both the nonsymbolic and symbolic ratio comparison tasks. We then tested at a group level whether children's correlation coefficients differed from zero with a one-sample t-test for each of the three heuristics on both the nonsymbolic and symbolic ratio tasks. In sum, a positive correlation between the Child's Deviation from Heuristic Model and the Ratio Deviation from Heuristic Model indicates choices based on a true ratio comparison when controlling for use of each heuristic, in contrast, a correlation of zero indicates use of the heuristic.

We also wanted to characterize which, if any, heuristics children used. To differentiate performing a ratio comparison from using a heuristic, we examined the trials where the ratio strategy predicted a different response from a given heuristic. Thus, we correlated the Ratio Deviation from Heuristic Model for each heuristic with a child's actual errors on the task (0 indicates a correct response, 1 indicates an incorrect response). A significant positive correlation indicates children made incorrect responses on trials where a given heuristic predicted an incorrect response. This analysis tests if the errors children made on the ratio comparison tasks were random or could be described by a bias toward using a specific heuristic. This analysis also produced six correlation coefficients for each subject, one for each heuristic for each task format. Again, we used a one-sample t-test to test whether the correlation differed from zero at a group level.

Mediation Analysis

We first removed any outliers on each task that were greater or less than three times the interquartile range. This process removed two ANS acuity scores. ANS acuity and symbolic ratio comparison scores were log-transformed to approach a normal distribution (ANS acuity Shapiro–Wilk $W = .96$; symbolic ratio comparison $W = .98$). To ensure that correlations between measures were not simply due to age or demographics of each school location, we partialed out age and school location from our measures of ANS acuity and symbolic and nonsymbolic ratio comparison, and school location from the age-standardized scores of the Woodcock–Johnson and the Key Math. The age and school location standardized Pearson correlations are reported in Table 1.

Mediation analyses test for a significant indirect effect (the product of the standardized coefficients a and b) that accounts for some portion of the original direct effect (c). The remaining direct effect is represented as c′. The goal of this analysis is to test whether nonsymbolic ratio comparison accuracy mediates the relation between ANS acuity and each symbolic math measure separately. In our cross-sectional experimental design this analysis cannot demonstrate causality. Instead, our goal is to examine whether nonsymbolic ratio calculation and ANS acuity account for the same or different variance in each symbolic math outcome measure. A significant mediation effect would be consistent with our hypothesis that nonsymbolic ratio calculation is a mechanism of the relation between ANS acuity and symbolic math.

Results

Symbolic and Nonsymbolic Ratio Comparison Performance

Children performed both the nonsymbolic (69% $t_{84} = 14.15$ $p < .001$, Cohen's $d = 1.54$) and symbolic (61% $t_{84} = 8.22$ $p < .001$, Cohen's $d = 0.89$) ratio tasks with above chance accuracy. This result held

Table 1
Age and School Location Standardized Pearson Correlations Between Each Measure

	ANS acuity	Nonsymbolic ratio comparison	Symbolic ratio comparison	Key-Math 3 Numeration	Key-Math 3 Data Analysis & Probability
Nonsymbolic ratio comparison	−.24*				
Symbolic ratio comparison	−.14	.52***			
Key-Math 3 Numeration	−.41***	.30**	.32**		
Key-Math 3 Data Analysis & Probability	−.28*	.37**	.30**	.64***	
W-J reading custer	−.18	.18	.30*	.28*	.23

*$p < .05$. **$p < .01$. ***$p < .001$.

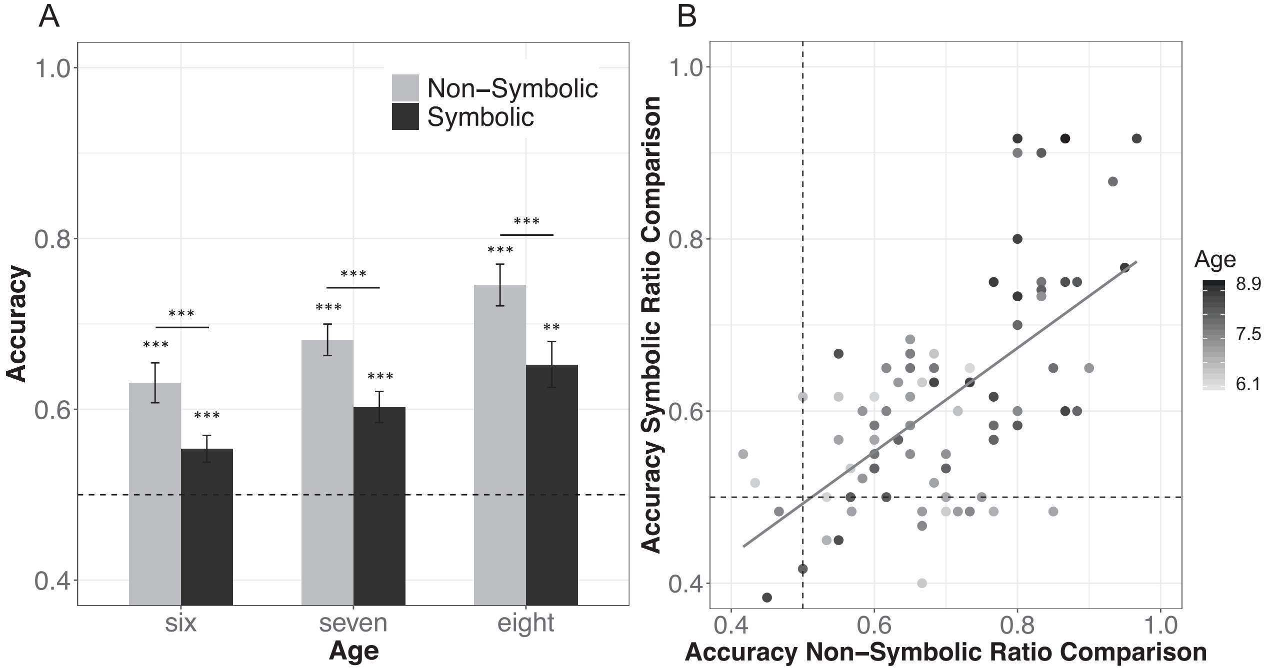

Figure 2. (A) Accuracy was above chance on both the nonsymbolic and symbolic ratio comparison tasks. Children were more accurate during nonsymbolic than symbolic ratio comparison. (B) Performance on the symbolic and nonsymbolic ratio comparison tasks were highly correlated. Dotted lines indicate chance performance. Error bars indicate the standard error of the mean. **$p < .01$, ***$p < .001$.

independently at each age tested (Figure 2A; Eight-year-olds Nonsymbolic 75% $t_{29} = 10.1$ $p < .001$, Cohen's $d = 1.84$; Symbolic 65% $t_{29} = 5.65$ $p < .001$, Cohen's $d = 1.03$; Seven-year-olds Nonsymbolic 68% $t_{32} = 9.8$ $p < .001$, Cohen's $d = 1.71$; Symbolic 60% $t_{32} = 5.62$ $p < .001$, Cohen's $d = 0.98$; Six-year-olds Nonsymbolic 63% $t_{21} = 5.61$ $p < .001$, Cohen's $d = 1.20$; Symbolic 55% $t_{21} = 3.43$ $p = .003$, Cohen's $d = 0.73$). Performance increased as a function of age on both ratio tasks (nonsymbolic $r_{83} = .42$, $t = 4.2$, $p < .001$; symbolic $r_{83} = .41$, $t = 4.1$, $p < .001$). Children were more accurate on nonsymbolic ratio comparison than symbolic ratio comparison ($t_{84} = 7.20$ $p < .001$, Cohen's $d = 0.78$). This format effect remained with the subset of children ($N = 59$) who recognized all numerals 1–30 on the numeral identification test, indicating that numeral recognition ability was not driving this difference ($t_{58} = 6.86$, $p < .001$, Cohen's $d = 0.89$). Moreover, this format effect was significant at each age independently (Eight-year-olds $t_{29} = 5.12$ $p < .001$, Cohen's $d = 0.93$; Seven-year-olds $t_{32} = 4.57$ $p < .001$, Cohen's $d = 0.80$; Six-year-olds $t_{21} = 2.80$ $p = .01$, Cohen's $d = 0.60$). As shown in Figure 2B, accuracy on the nonsymbolic and symbolic ratio comparison tasks were highly correlated (Table A1 in Appendix S1; $r = .62$, $t_{83} = 7.22$ $p < .001$).

To test whether the ratio of ratios of the two arrays impacted accuracy, we ran a generalized mixed-effects linear model following a binomial distribution predicting accuracy as a function of task format (symbolic or nonsymbolic), the ratio of ratios as fixed effects, and a random effect of subject. This model indicated significant main effects of ratio and task (ratio of ratio $\beta = .11$, $z = 9.46$, $p < .001$; task $\beta = -.16$, $z = -2.21$, $p = .03$), and an interaction between ratio of ratio and task format ($\beta = -.06$, $z = -3.90$, $p < .001$). Follow-up analyses indicate that accuracy was dependent on the ratio of ratios for both formats of the ratio comparison task when modeled separately (Nonsymbolic ratio of ratio $\beta = .11$, $z = 9.52$, $p < .001$; Symbolic $\beta = .05$, $z = 4.92$, $p < .001$).

Nonsymbolic Ratio Comparison Facilitates Symbolic Ratio Comparison

To test the hypothesis that experience with nonsymbolic ratio comparison scaffolds understanding of the symbolic version of the task, we ran a mixed-effects analysis of variance (ANOVA) with a main effect of task order (first or second), task format (symbolic or nonsymbolic), a task order by task format interaction, and a random effect of subject. As planned, we removed any children who could not identify all numerals 1–30 from these analyses (see Figure A4 in Appendix S1 for the same analysis with all subjects). This left 59 children spanning the age

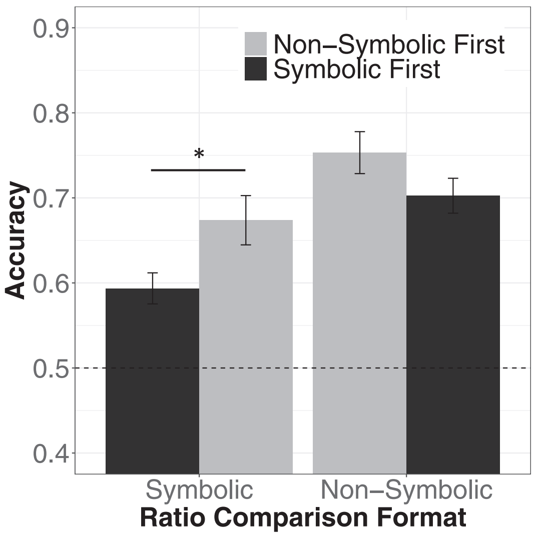

Figure 3. There is a significant interaction between task order and task format. Children who completed the nonsymbolic ratio comparison task first were significantly more accurate on the symbolic ratio comparison task. Error bars indicate the standard error of the mean. $*p < .05$.

range of the entire sample (11 six-year-olds, 22 seven-year-olds, 26 eight-year-olds). In line with our prediction, there was a significant task order by task format interaction (Figure 3; $F_{1,57} = 5.03$, $p = .03$). This effect was driven by significantly better performance on the symbolic ratio comparison test after completion of the nonsymbolic version of the task (59% vs. 67% $t_{57} = -2.44$, $p = .02$, Cohen's $d = 0.64$) and a nonsignificant effect of task order on nonsymbolic ratio comparison accuracy (70% vs. 75% $t_{57} = -1.59$, $p = .12$). The ANOVA also revealed a significant main effect of task (72% vs. 63% $F_{1,57} = 47.1$, $p < .001$), consistent with our analysis of the larger data set, and a nonsignificant effect of task order (First 67% vs. Second 68% $F_{1,57} = 1.09$, $p = .30$). These findings are consistent with our hypothesis that nonsymbolic ratio reasoning serves as a scaffold for symbolic ratio reasoning.

Children Perform a True Ratio Comparison But Also Use Incorrect Heuristics

We examined whether children performed a ratio comparison when controlling for heuristic use on both the symbolic and nonsymbolic ratio tasks. We included the 82 subjects in this analysis who completed all trials. Figure 4 displays the correlation between each Ratio Deviation from Heuristic Model and Child's Deviation from Heuristic Model. A one-sample t-test revealed a significant positive correlation (Figure 4; Nonsymbolic More Good $r = .35$, $t_{81} = 11.6$, $p < .001$; More Items $r = .41$, $t_{81} = 15.1$, $p < .001$; Less Bad $r = .44$, $t_{81} = 16.8$, $p < .001$; Symbolic More Good $r = .18$, $t_{81} = 6.49$, $p < .001$; More Items $r = .23$, $t_{81} = 8.42$, $p < .001$; Less Bad $r = .25$, $t_{81} = 9.28$, $p < .001$). This analysis revealed that children's choices were significantly driven by a comparison of ratios in both the symbolic and nonsymbolic ratio comparison tasks even when controlling for the potential use of each incorrect heuristic.

Although the previous analysis demonstrates that children did not rely on any particular unidimensional heuristic to solve the ratio comparison tasks, it remains possible that children used one or more heuristic on some trials. Figure 5 displays the correlation between the Ratio Deviation from Heuristic Model for each strategy and a child's actual errors on the nonsymbolic and symbolic ratio comparison tasks (for a breakdown of positive correlations by subject see Figure A3 in Appendix S1) For the nonsymbolic ratio comparison task, a one-sample t-test on the mean correlation indicated that children's errors were in accordance with the More Good or More Items strategies (Figure 5; More Good $r = .36$, $t_{81} = 10.7$, $p < .001$; More Items $r = .29$, $t_{81} = 9.12$, $p < .001$). The negative correlation between errors and the Less Bad strategy indicated that children were significantly doing the opposite of the Less Bad strategy ($r = -.24$, $t_{81} = -8.11$, $p < .001$). The same pattern of results was found for the symbolic ratio comparison task. Specifically, children's errors were in accord with the More Good ($r = .25$, $t_{81} = 7.83$, $p < .001$) and the More Items ($r = .20$, $t_{81} = 6.61$, $p < .001$) strategy. Children also did the opposite of the Less Bad strategy on the symbolic version of the task ($r = -.17$, $t_{81} = -6.17$, $p < .001$). Thus, the errors children made on the ratio task were qualitatively similar regardless of presentation format.

No Evidence of a Strategy Shift After Nonsymbolic Ratio Practice

We used the same strategy analysis described above to investigate heuristic use, but this time split participants by the order in which they completed the ratio tasks (nonsymbolic first $n = 27$, symbolic first $n = 36$) and analyzed their performance on only the symbolic version of the task. Contrary to our hypothesis, there was no significant difference in the degree to which children's errors reflected use of the More Good or More Items strategies by the order in which they completed the symbolic

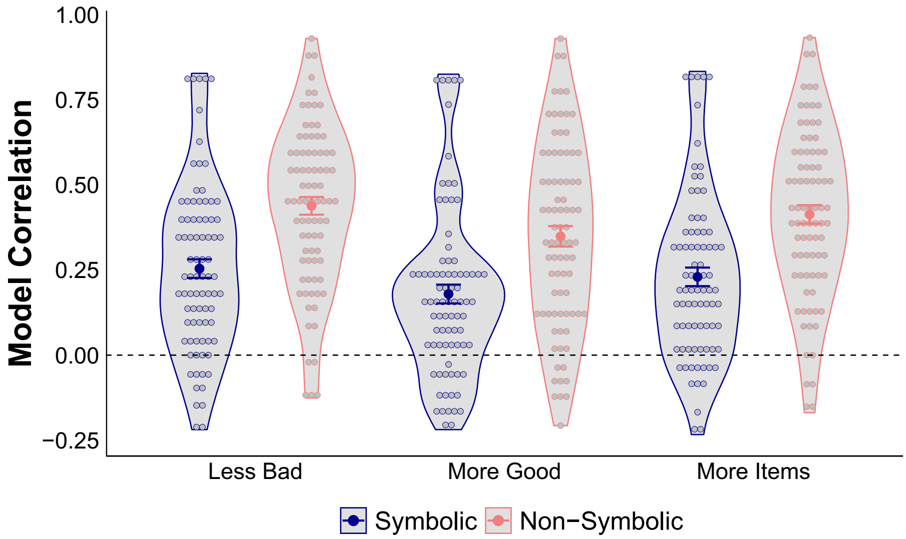

Figure 4. Correlation between each Ratio Deviation from Heuristic Model and Child's Deviation from Heuristic Model. A correlation significantly above zero indicates children made choices consistent with ratio comparison, even when controlling for the use of each incorrect heuristic. Children's responses were significantly positively correlated with the responses indicated by a ratio comparison strategy for all three heuristics on both formats of the ratio comparison task $p < .001$. Points indicate the mean correlation of all subjects, and the gray area of the violin plot indicates the distribution of individual correlations. Error bars indicate the standard error of the mean. [Color figure can be viewed at wileyonlinelibrary.com]

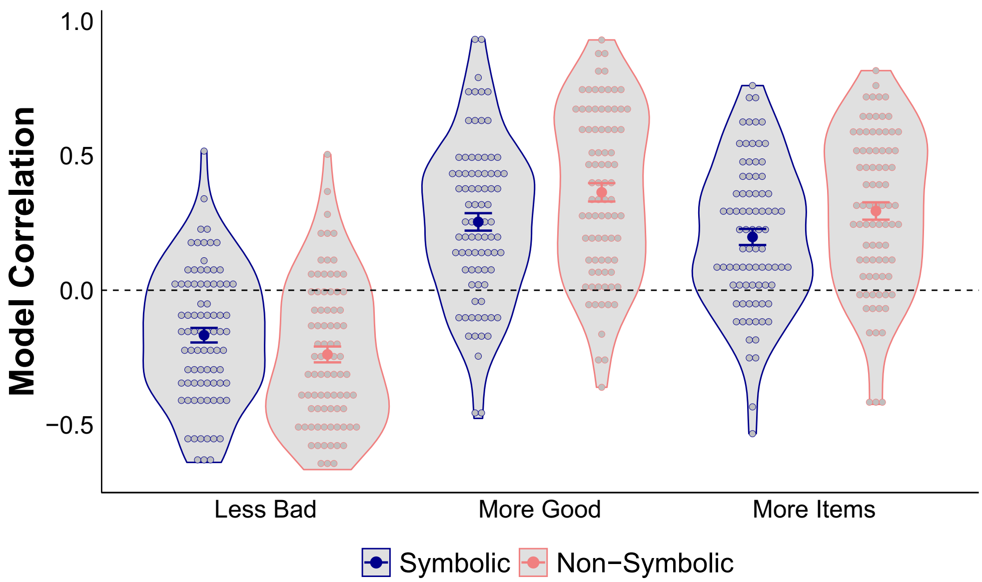

Figure 5. Correlation between the Ratio Deviation from Heuristic Model for each strategy and a child's actual errors on the nonsymbolic and symbolic ratio comparison tasks. A correlation significantly above zero indicates the use of the strategy. The errors children made on both the symbolic and nonsymbolic ratio tasks were consistent with the use of the More Good or More Items strategy $p < .001$. The errors children made were significantly opposite of using the Less Bad strategy $p < .001$. Points indicate the mean correlation of all subjects, and the gray area of the violin plot indicates the distribution of individual correlations. Error bars indicate the standard error of the mean. [Color figure can be viewed at wileyonlinelibrary.com]

task (Figure A1 in Appendix S1; two-sample t test More Good $t_{61} = -.66$, $p = .51$; More Items $t_{61} = -.82$, $p = .41$). There was also no effect of task order on the degree to which children did not use the Less Bad heuristic (Figure A1 in Appendix S1; $t_{61} = 1.09$, $p = .28$).

The Relation Between ANS Acuity, Nonsymbolic Ratio Comparison, and Symbolic Math

Finally, we examined how children's ability to compare ratios nonsymbolically related to their ANS acuity, and symbolic math skill. Table 1 displays the age and school location standardized correlations between nonsymbolic and symbolic ratio comparison skill, ANS acuity, Key-Math-3 Numeration and Data Analysis and Probability sections, and the Reading Cluster score on the Woodcock–Johnson (for zero-order correlations see Table A1 in Appendix S1). Mean performance on the Key-Math-3 Numeration subtest was 9.05 (*SD* 3.3), and mean performance on the Data Analysis & Probability subtest was 8.47 (*SD* 3.1). The scale scores are constructed to have a mean of 10 (*SD* 3).

Children completed three symbolic math measures: symbolic ratio comparison, Key-Math-3 Numeration, and Key-Math-3 Data Analysis and Probability. ANS acuity was no longer correlated with accuracy on the symbolic ratio comparison task when controlling for age and school location (comparison of Table A1 in Appendix S1 and Table 1). Consequently, we ran mediation models with only the Key-Math 3 Numeration and Data Analysis and Probability sections as symbolic math outcome measures. We ran all mediation analyses using the "mediation" package in R (Tingle, Yamamoto, Hirose, Keele & Imai, 2014). As seen in Table 1, the Key-Math-3 Numeration section score was significantly correlated with children's score on the Woodcock–Johnson Basic Reading Skills. To ensure that this outcome measure is a test of general math skill, and not general academic performance, we partialed out children's Woodcock–Johnson Basic Reading score from the Key-Math Numeration score. The residuals from the models with age, school location, and the Basic Reading Skills test as predictors were used in the mediation analysis.

We tested whether nonsymbolic ratio comparison accuracy mediated the relation between ANS acuity and the two subtests of the Key-Math-3 test (Figure 6). We first ran the mediation analysis with the Numeration subtest as an outcome measure. ANS acuity was a significant predictor of the Key-Math 3 Numeration score (standardized $\beta = -.41$, $p < .001$) and was also a significant predictor of accuracy on the nonsymbolic ratio comparison task (standardized $\beta = -.27$, $p = .02$). However, when ANS acuity and nonsymbolic ratio comparison accuracy were both entered into the same model, nonsymbolic ratio comparison accuracy was no

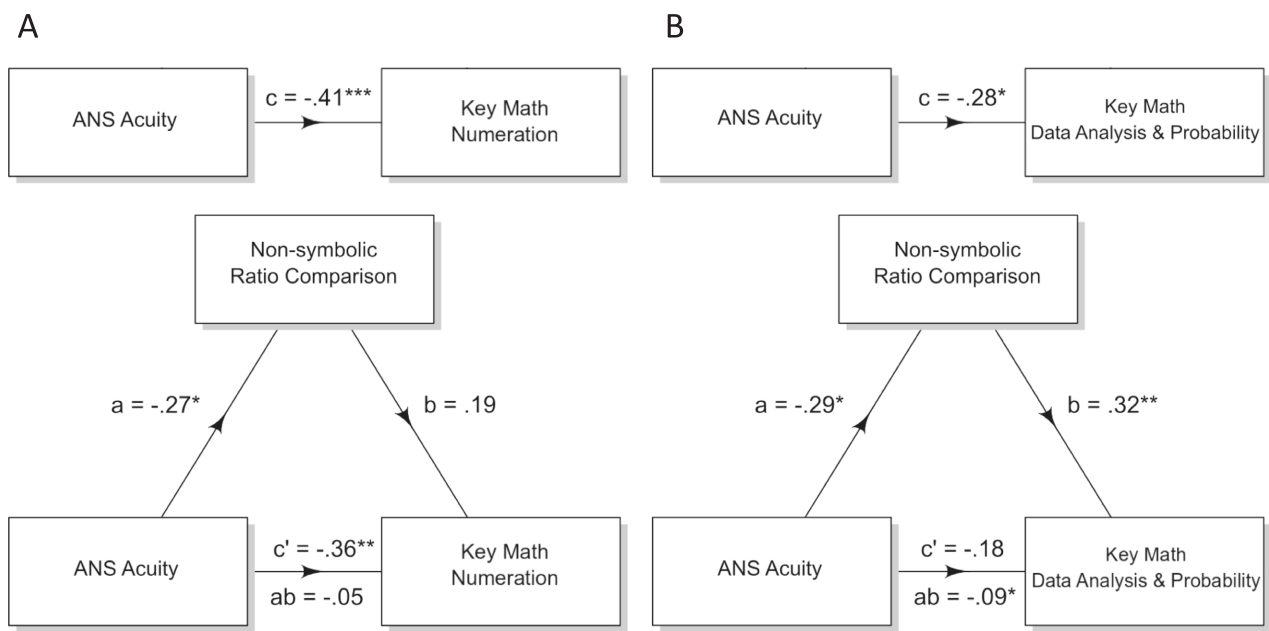

Figure 6. Mediation analyses test for a significant indirect effect (the product of the standardized coefficients a and b) that accounts for some portion of the original direct effect (c). The remaining direct effect is represented as c′. (A) Nonsymbolic ratio comparison accuracy does not mediate the relation between Approximate Number System (ANS) acuity and a child's score on the Key-Math-3 Numeration section. The direct effect c′ remains significant. (B) Nonsymbolic ratio comparison accuracy mediates the relation between ANS acuity and a child's score on the Key-Math-3 Data Analysis and Probability section. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

longer a significant predictor of the Key-Math-3 Numeration score (β = .19, p = .09), whereas ANS acuity remained a significant predictor (β = −.36, p = .002). The direct effect was significant when tested with a bootstrap estimation approach with 5000 simulations (direct effect = −.36, 95% CI [−.54, −.18], p < .001), whereas the indirect effect was not significant (indirect effect = −.05, 95% CI [−.15, .001], p = .06). Thus, nonsymbolic ratio comparison did not mediate the relation between ANS acuity and a child's score on the Key-Math-3 Numeration test. Instead, ANS acuity explained unique variance beyond nonsymbolic ratio calculation accuracy in performance on the Key-Math-3 Numeration test.

We then ran the mediation analysis with the Data Analysis and Probability subtest as an outcome measure. ANS acuity was a significant predictor of a child's score on the Key-Math-3 Data Analysis and Probability section (standardized β = −.28, p = .03) and of accuracy on the nonsymbolic ratio comparison task (standardized β = −.29, p = .02). In contrast to the previous analysis of the Numeration subtest, ANS acuity was no longer a significant predictor of the score on the Probability and Data Analysis subtest after controlling for the mediator, nonsymbolic ratio comparison accuracy (ANS acuity standardized β = −.18, p = .13; nonsymbolic ratio comparison accuracy standardized β = .32, p = .008). Nonsymbolic ratio comparison mediates the relation between ANS acuity and performance on the Key-Math-3 Data Analysis and Probability subtest. The indirect effect was significant when tested with a bootstrap estimation approach (indirect effect = −.09, 95% CI [−.22, −.01], p = .03). The direct effect was not significant (direct effect = −.18, 95% CI [−.40, .02], p = .07). The proportion mediated was .34 (p = .03, 95% CI [0.04, 1.10]). Thus, a higher Key-Math-3 Data Analysis and Probability score was associated with 0.09 *SD*s sharper ANS acuity as mediated through nonsymbolic ratio comparison accuracy.

Discussion

The first goal of our study was to assess whether elementary school children engage in both nonsymbolic and symbolic ratio reasoning, and whether exposure to nonsymbolic ratio reasoning might scaffold symbolic ratio reasoning. We found that children are adept at solving both nonsymbolic and symbolic ratio comparisons before they receive formal education about fractions and ratios. While performance on both tasks improved with age, even 6-year-olds demonstrated above chance accuracy on both the nonsymbolic and symbolic ratio tasks. Our results support previous work demonstrating that young children engage in intuitive ratio reasoning before they learn about ratios and fractions in school (Boyer & Levine, 2015; Falk et al., 2012; Yost et al., 1962). We further provide the first evidence that young children intuitively reason about ratios with symbolic numerals. While one previous study found that slightly older 7- to 10-year-old children could compare ratios presented as a frequency distribution (e.g., 4 tokens out of 10 are gold), the stimuli in that study were constructed such that children could have attended only to the numerator rather than truly evaluating the ratio of the numerator to the denominator (Ruggeri, Vagharchakian, & Xu, 2018). This study is thus the first to demonstrate that 6- to 8-year-old children can perform approximate ratio comparisons with a symbolic format without relying on unidimensional comparison heuristics.

Consistent with our hypothesis that engaging in nonsymbolic ratio reasoning scaffolds symbolic ratio reasoning, we found a significant effect of task order on symbolic ratio comparison accuracy. Children who received the nonsymbolic task first performed with higher accuracy on the symbolic task as compared to children who started the session with the symbolic task. One possibility is that engaging in the nonsymbolic task first highlighted the conceptual link between the two tasks. The nonsymbolic task offers a concrete representation of the ratio comparison problem, which may allow children to create a conceptual model of correct ratio comparison. This conceptual model may remain opaque when children encounter only symbolic representations of the numerical magnitudes involved in the computation. This interpretation is consistent with prior work where nonsymbolic representations of a math problem boost symbolic calculation (for examples see Carbonneau, Marley, & Selig, 2013; Fujimura, 2001; Fyfe, McNeil, Son, & Goldstone, 2014; Park et al., 2016) An alternative possibility is that because overall performance is greater on the nonsymbolic task, engaging in the nonsymbolic ratio comparison task first boosted children's domain-general or domain-specific confidence. Under this scenario, practice on the nonsymbolic task may not have changed how children solved the symbolic task, but instead increased confidence in their ratio intuitions. Our finding that children did not change the strategies they used on the symbolic task after completion of the nonsymbolic task, but still had an overall increase in accuracy, may be interpreted as support for this hypothesis. This type of confidence hysteresis effect, where children perform with higher accuracy when trials move from easy to hard, has been demonstrated in other domains and among ANS tasks (Hock & Schöner, 2010; Odic, Hock, & Halberda, 2014). Ultimately, future work utilizing a pretest-training-posttest experimental design is necessary to disentangle the mechanism of this scaffolding effect and to test the robustness of the benefit of nonsymbolic ratio practice for symbolic ratio understanding.

The second goal of our experiment was to examine whether children use unidimensional heuristics when solving the ratio comparison tasks. Importantly, children's performance on the two ratio tasks could not be explained by any single incorrect heuristic. At the same time, when children did make an error, they tended to erroneously pick the machine with more of the preferred item, or with more total items. This finding is consistent with previous research on young children's intuitive ratio reasoning (Clarke & Roche, 2009; Falk et al., 2012; Jeong et al., 2007; Obersteiner et al., 2015; Shaklee & Paszek, 1985; Siegler et al., 1981). Taken together, our results suggest that children are capable of true ratio reasoning, but continue to exhibit a bias toward the ratio with the greater number of preferred items and the ratio with a greater number of total items. Children also exhibit this pattern of mixed strategy use when placing a fraction magnitude on a number line (Braithwaite & Siegler, 2018). Future work with an even larger stimulus set should examine whether children's use of incorrect unidimensional heuristics depends on the difficulty of the ratios being compared. Perhaps children reply on incorrect heuristics when the ratio comparison is difficult.

The third and final goal of this study was to examine whether the ANS serves as a cognitive foundation for nonsymbolic ratio reasoning, and to test whether performing a nonsymbolic operation may be a mechanism of the relation between ANS acuity and symbolic math. Our hypothesis was that sharper ANS acuity would allow children to create better intuitive ratio models, which allow for more accurate nonsymbolic ratio calculation. In turn, better nonsymbolic ratio comparison accuracy can provide a better conceptual scaffold for related math concepts. Consistent with the first part of our hypothesis, ANS acuity was significantly correlated with children's accuracy on the nonsymbolic ratio reasoning task. Moreover, performance on both the nonsymbolic and symbolic ratio reasoning tasks was dependent on the ratio of ratios being compared, in accord with *Weber's* law (Feigenson et al., 2004). Consistent with the second part of our hypothesis, nonsymbolic ratio accuracy mediated the relation between ANS acuity and performance on the Key-Math-3 Data Analysis and Probability section. This Key-Math-3 section included questions such as "Which spinner gives you an equal chance of landing on green or white?" and "There are two more turtles than which animal?" while displaying a pictograph. Thus, nonsymbolic ratio comparison mediates the relation between ANS acuity and formal math questions about probability and counting objects. This finding is in line with work demonstrating that performing a nonsymbolic operation explains significant unique variance in formal math skill beyond ANS acuity (Matthews et al., 2016; Park & Brannon, 2014; Pinheiro-Chagas et al., 2014; Starr et al., 2016).

Our finding that nonsymbolic ratio comparison explains significant unique variance in performance on the Key-Math-3 Data Analysis and Probability subtest is consistent with the finding that nonsymbolic, approximate ratio reasoning predicts symbolic fraction, number line, and algebra knowledge when controlling for ANS acuity in adults (Matthews et al., 2016). These findings have led to the proposal of an intuitive Ratio Processing System (Matthews & Hubbard, 2017). The current experiment is the first to find similar relations between nonsymbolic, approximate ratio processing ability, ANS acuity, and formal math skill in elementary school children. Thus, the Ratio Processing System may support children's early understanding of ratios during the elementary school years, even without explicit instruction linking children's nonsymbolic and symbolic ratio representations.

However, we did not find evidence that nonsymbolic ratio accuracy mediates the relation between ANS acuity and scores on the Key-Math-3 Numeration section. In this case, the direct relation between ANS acuity and basic numeration concepts was not significantly attenuated by the indirect effect of nonsymbolic ratio accuracy. This finding is consistent with work correlating sharper ANS acuity with better early symbolic math skills in children (Schneider et al., 2016). The Key-Math-3 Numeration section included questions such as "Read these numbers from least to greatest" and "Starting at forty-one, count up by tens". Thus, the Key-Math-3 Numeration section tests general symbolic number skills, whereas the Data Analysis & Probability section tests concepts that are more similar to the nonsymbolic ratio task. Taken together, our mediation analyses suggest that nonsymbolic ratio calculations serve as a conceptual scaffold for symbolic math problems that conceptually overlap, but not for unrelated math skills. In other words, perhaps not surprisingly, nonsymbolic calculation cannot function as a conceptual scaffold for symbolic calculation when the required calculations differ dramatically from each other. While the current experiment explored the relations between ANS acuity, a nonsymbolic ratio comparison operation, and symbolic math, future research should test whether a more general measure of nonsymbolic approximate calculation, beyond ratio calculation alone, mediates the relation between ANS acuity and general math skill. Such work would identify the degree to which algorithmic overlap between the nonsymbolic and symbolic math tested is required for mediation of the relation between ANS acuity and symbolic math.

While we initially predicted that nonsymbolic ratio reasoning would mediate the relation between ANS acuity and approximate, symbolic ratio reasoning performance, we were unable to test this hypothesis because performance on our symbolic ratio reasoning task was no longer correlated with ANS acuity after controlling for age and school location. This lack of correlation may be due to the variability in numeral knowledge between schools. For example, perhaps numeral identification (especially numbers 11–30) was taught at some schools and not others. In support of this hypothesis, two schools were significant predictors of numeral identification score in a model with age and school location as regressors (Table A3 in Appendix S1). In contrast to our standardized Key-Math-3 measures, curriculum timing may have had a role to play in our measurement of symbolic ratio comparison ability, and this may have made it harder to measure the correlation between symbolic ratio comparison ability and ANS acuity.

It is our hope that the current experiment will inspire future explorations of how nonsymbolic numerical reasoning can scaffold symbolic numerical reasoning. A limitation of this study is that both ratio tasks were presented in a part-part format rather than the part-whole format used with symbolic fractions (e.g. X out of X gumballs are blue). For example, it is possible we could see an increased use of the More Items strategy if the total number of items was explicitly displayed during the symbolic version of the task. The total number of gumballs in the nonsymbolic task may be more directly perceived as a whole by ignoring the individual gumball color. Future work could explore whether presenting the task in a part-whole format changes heuristic use or influences the degree of format facilitation. Future work should also explore the role of executive function and short-term memory in the relations between nonsymbolic ratio reasoning, ANS acuity, and symbolic math given the importance of these skills for general mathematics.

In summary, our findings highlight that even before children begin to learn about fractions and ratios in school, they possess an intuitive ratio reasoning capacity. This intuitive ratio reasoning skill extends to nonsymbolic quantities and numerals. Children have a notoriously difficult time understanding ratios in school (Siegler, Fazio, Bailey, & Zhou, 2013). This difficulty has sometimes been attributed to a "Whole Number Bias", the tendency for children to overgeneralize integer calculation to rational numbers (DeWolf & Vosniadou, 2015; Ni & Zhou, 2005). Emphasizing children's intuitions about ratios as well as whole numbers at an early age holds the potential to mitigate some of this conceptual confusion. Specifically, grounding children's symbolic ratio reasoning in isomorphic nonsymbolic problems may facilitate learning. A greater understanding of the intuitive mathematical knowledge children possess may help educators make abstract mathematical computations more accessible to young learners.

References

Ahl, V. A., Moore, C. F., & Dixon, J. A. (1992). Development of intuitive and numerical proportional reasoning. *Cognitive Development*, *7*, 81–108. https://doi.org/10.1016/0885-2014(92)90006-D

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition*, *98*, 199–222. https://doi.org/10.1016/j.cognition.2004.09.011

Boyer, T. W., & Levine, S. C. (2015). Prompting children to reason proportionally: Processing discrete units as continuous amounts. *Developmental Psychology*, *51*, 615–620. https://doi.org/10.1037/a0039010

Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: Where young children go wrong. *Developmental Psychology*, *44*, 1478–1490. https://doi.org/10.1037/a0013110

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Braithwaite, D. W., & Siegler, R. S. (2018). Developmental changes in the whole number bias. *Developmental Science*, *21*, e12541. https://doi.org/10.1111/desc.12541

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From brain to education. *Science*, *332*, 1049–1053. https://doi.org/10.1126/science.1201536

Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics

with concrete manipulatives. *Journal of Educational Psychology*, *105*, 380–400. https://doi.org/10.1037/a0031084
Clarke, D. M., & Roche, A. (2009). Students' fraction comparison strategies as a window into robust understanding and possible pointers for instruction. *Educational Studies in Mathematics*, *72*, 127–138. https://doi.org/10.1007/s10649-009-9198-9
Connolly, A. J. (2007). *KeyMath-3 diagnostic assessment: Manual forms A and B*. Minneapolis, MN: Pearson.
Denison, S., & Xu, F. (2014). The origins of probabilistic inference in human infants. *Cognition*, *130*, 335–347. https://doi.org/10.1016/j.cognition.2013.12.001
DeWind, N. K., Adams, G. K., Platt, M. L., & Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition*, *142*, 247–265. https://doi.org/10.1016/j.cognition.2015.05.016
DeWolf, M., & Vosniadou, S. (2015). The representation of fraction magnitudes and the whole number bias reconsidered. *Learning and Instruction*, *37*, 39–49. https://doi.org/10.1016/j.learninstruc.2014.07.002
Drucker, C. B., Rossa, M. A., & Brannon, E. M. (2015). Comparison of discrete ratios by rhesus macaques (*Macaca mulatta*). *Animal Cognition*, *19*, 75–89. https://doi.org/10.1007/s10071-015-0914-9
Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428
Eckert, J., Call, J., Hermes, J., Herrmann, E., & Rakoczy, H. (2018). Intuitive statistical inferences in chimpanzees and humans follow Weber's law. *Cognition*, *180*, 99–107. https://doi.org/10.1016/j.cognition.2018.07.004
Falk, R., & Wilkening, F. (1998). Children's construction of fair chances: Adjusting probabilities. *Developmental Psychology*, *34*, 1340. https://doi.org/10.1037/0012-1649.34.6.1340
Falk, R., Yudilevich-Assouline, P., & Elstein, A. (2012). Children's concept of probability as inferred from their binary choices—Revisited. *Educational Studies in Mathematics*, *81*, 207–233. https://doi.org/10.1007/s10649-012-9402-1
Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314. https://doi.org/10.1016/j.tics.2004.05.002
Fujimura, N. (2001). Facilitating children's proportional reasoning: A model of reasoning processes and effects of intervention on strategy change. *Journal of Educational Psychology*, *93*, 589–603. https://doi.org/10.1037//0022-0663.93.3.589
Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review*, *26*, 9–25. https://doi.org/10.1007/s10648-014-9249-3
Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, *447*, 589–591. https://doi.org/10.1038/nature05850
Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature Letters*, *455*, 665–668. https://doi.org/10.1038/nature07246
Hock, H., & Schöner, G. (2010). Measuring perceptual hysteresis with the modified method of limits: Dynamics at the threshold. *Seeing and Perceiving*, *23*, 173–195. https://doi.org/10.1163/187847510X503597
Honoré, N., & Noël, M.-P. (2016). Improving preschoolers' arithmetic through number magnitude training: The impact of non-symbolic and symbolic training. *PLoS One*, *11*, e0166685. https://doi.org/10.1371/journal.pone.0166685
Jeong, Y., Levine, S. C., & Huttenlocher, J. (2007). The development of proportional reasoning: Effect of continuous versus discrete quantities. *Journal of Cognition and Development*, *8*, 237–256. https://doi.org/10.1080/15248370701202471
Jordan, N. C., Resnick, I., Rodrigues, J., Hansen, N., & Dyson, N. (2016). Delaware longitudinal study of fraction learning implications for helping children with mathematics difficulties. *Journal of Learning Disabilities*, *50*, 621–630. https://doi.org/10.1177/0022219416662033
Kayhan, E., Gredebäck, G., & Lindskog, M. (2017). Infants distinguish between two events based on their relative likelihood. *Child Development*, *89*, e507–e519. https://doi.org/10.1111/cdev.12970
Kibbe, M. M., & Feigenson, L. (2017). A dissociation between small and large numbers in young children's ability to "solve for x" in non-symbolic math problems. *Cognition*, *160*, 82–90. https://doi.org/10.1016/j.cognition.2016.12.006
Kleiner, M., Brainard, D., & Pelli, D. (2007). "What's new in Psychtoolbox-3?" Perception 36 ECVP Abstract Supplement.
Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schönmann, C., Plangger, F., . . . von Aster, M. (2011). Mental number line training in children with developmental dyscalculia. *NeuroImage*, *57*, 782–795. https://doi.org/10.1016/j.neuroimage.2011.01.070
Lewis, M. R., Matthews, P. G., & Hubbard, E. M. (2016). Neurocognitive architectures and the nonsymbolic foundations of fractions understanding. In *Development of mathematical cognition* (pp. 141–164). Elsevier. https://doi.org/10.1016/B978-0-12-801871-2.00006-X
Lindskog, M., Winman, A., Juslin, P., & Poom, L. (2013). Measuring acuity of the approximate number system reliably and validly: The evaluation of an adaptive test procedure. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00510
Matthews, P. G., & Hubbard, E. M. (2017). Making space for spatial proportions. *Journal of Learning Disabilities*, *50*, 644–647. https://doi.org/10.1177/0022219416679133
Matthews, P. G., Lewis, M. R., & Hubbard, E. M. (2016). Individual differences in nonsymbolic ratio processing predict symbolic math performance. *Psychological Science*, *27*, 191–202. https://doi.org/10.1177/0956797615617799

McCrink, K., Shafto, P., & Barth, H. (2016). The relationship between non-symbolic multiplication and division in childhood. *The Quarterly Journal of Experimental Psychology*, *70*, 686–702. https://doi.org/10.1080/17470218.2016.1151060

McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science*, *18*, 740–745. https://doi.org/10.1111/j.1467-9280.2007.01969.x

Möhring, W., Newcombe, N. S., Levine, S. C., & Frick, A. (2016). Spatial proportional reasoning is associated with formal knowledge about fractions. *Journal of Cognition and Development*, *17*, 67–84. https://doi.org/10.1080/15248372.2014.996289

National Center for Education Statistics. (2017). National Assessment of Educational Progress: An overview of NAEP. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education.

Ni, Y., & Zhou, Y.-D. (2005). Teaching and learning fraction and rational numbers: the origins and implications of whole number bias. *Educational Psychologist*, *40*(1), 27–52. https://doi.org/10.1207/s15326985ep4001_3

Obersteiner, A., Bernhard, M., & Reiss, K. (2015). Primary school children's strategies in solving contingency table problems: The role of intuition and inhibition. *ZDM*, *47*, 825–836. https://doi.org/10.1007/s11858-015-0681-8

Odic, D., Hock, H., & Halberda, J. (2014). Hysteresis affects approximate number discrimination in young children. *Journal of Experimental Psychology: General*, *143*, 255–265. https://doi.org/10.1037/a0030825

Park, J., Bermudez, V., Roberts, R. C., & Brannon, E. M. (2016). Non-symbolic approximate arithmetic training improves math performance in preschoolers. *Journal of Experimental Child Psychology*, *152*, 278–293. https://doi.org/10.1016/j.jecp.2016.07.011

Park, J., & Brannon, E. M. (2014). Improving arithmetic performance with number sense training: An investigation of underlying mechanism. *Cognition*, *133*, 188–200. https://doi.org/10.1016/j.cognition.2014.06.011

Parsons, S., & Bynner, J. (2005). *Does numeracy matter more? National Research and Development Centre for Adult Literacy and Numeracy*. London, UK: The Basic Skills Agency.

Pelli, D. G. (1997). The video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Pinheiro-Chagas, P., Wood, G., Knops, A., Krinzinger, H., Lonnemann, J., Starling-Alves, I., . . . Haase, V. G. (2014). In How many ways is the approximate number system associated with exact calculation? *PLoS One*, *9*, e111155. https://doi.org/10.1371/journal.pone.0111155

Rakoczy, H., Clüver, A., Saucke, L., Stoffregen, N., Gräbener, A., Migura, J., & Call, J. (2014). Apes are intuitive statisticians. *Cognition*, *131*, 60–68. https://doi.org/10.1016/j.cognition.2013.12.011

Ruggeri, A., Vagharchakian, L., & Xu, F. (2018). Icon arrays help younger children's proportional reasoning. *British Journal of Developmental Psychology*. https://doi.org/10.1111/bjdp.12233

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2016). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*, e12372. https://doi.org/10.1111/desc.12372

Shaklee, H., & Paszek, D. (1985). Covariation judgment: Systematic rule use in middle childhood. *Child Development*, *56*, 1229–1240. https://doi.org/10.2307/1130238

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., . . . Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*, 691–697. https://doi.org/10.1177/0956797612440101

Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in Cognitive Sciences*, *17* (1), 13–19. https://doi.org/10.1016/j.tics.2012.11.004

Siegler, R. S., Strauss, S., & Levin, I. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child development*, *46*, 1–84.

Starr, A., Roberts, R. C., & Brannon, E. M. (2016). Two potential mechanisms underlying the link between approximate number representations and symbolic math in preschool children. *Proceedings of the Cognitive Science Society*. Retrieved from https://www.researchgate.net/profile/Ariel_Starr/publication/308699729_Two_Potential_Mechanisms_Underlying_the_Link_between_Approximate_Number_Representations_and_Symbolic_Math_in_Preschool_Children/links/57ebf30408ae92eb4d263dec.pdf

Szkudlarek, E., & Brannon, E. M. (2018). Approximate arithmetic training improves informal math performance in low achieving preschoolers. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00606

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software*, *59*, 1–38. http://www.jstatsoft.org/v59/i05/

Vallentin, D., & Nieder, A. (2008). Behavioral and prefrontal representation of spatial proportions in the monkey. *Current Biology*, *18*, 1420–1425. https://doi.org/10.1016/j.cub.2008.08.042

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 5012–5015. https://doi.org/10.1073/pnas.0704450105

Yost, P. A., Siegel, A. E., & Andrews, J. M. (1962). Nonverbal probability judgments by young children. *Child Development*, *33*, 769. https://doi.org/10.2307/1126888

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Appendix S1. Supplementary material