Journal of Experimental Child Psychology 214 (2022) 105277

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Children's gesture use provides insight into proportional reasoning strategies

Michelle A. Hurst [a,*], Alyson Wong [b], Raychel Gordon [b,1], Aziza Alam [b,2], Sara Cordes [b]

[a] Department of Psychology, University of Chicago, Chicago, IL 60615, USA
[b] Department of Psychology, Boston College, Chestnut Hill, MA 02467, USA

ARTICLE INFO

Article history:
Received 7 October 2020
Revised 27 July 2021
Available online 6 September 2021

Keywords:
Proportional reasoning
Gesture
Numerical interference
Whole number bias
Continuous proporiton
Discrete proportion
Equivalent proportion

ABSTRACT

Children struggle with proportional reasoning when discrete countable information is available because they over-rely on this numerical information even when it leads to errors. In the current study, we investigated whether different types of gesture can exacerbate or mitigate these errors. Children aged 5–7 years (N = 135) were introduced to equivalent proportions using discrete gestures that highlighted separate parts, continuous gestures that highlighted continuous amounts, or no gesture. After training, children completed a proportional reasoning match-to-sample task where whole number information was occasionally pitted against proportional information. After the task, we measured children's own gesture use. Overall, we did not find condition differences in proportional reasoning; however, children who observed continuous gestures produced more continuous gestures than those who observed discrete gestures (and vice versa for discrete gestures). Moreover, producing fewer discrete gestures and more continuous gestures was associated with lower numerical interference on the match-to-sample task. Lastly, to further investigate individual differences, we found that children's inhibitory control and formal math knowledge were correlated with proportional reasoning in general but not with numerical interference in particular. Taken together, these findings highlight that children's own gestures

* Corresponding author.
E-mail address: hurstm@uchicago.edu (M.A. Hurst).
1 Current address: College of Education, University of Maryland, College Park, MD 20742, USA.
2 Current address: Teachers College, Columbia University, New York, NY 10027, USA.

https://doi.org/10.1016/j.jecp.2021.105277

may be a powerful window into the information they attend to during proportional reasoning.

Introduction

Infants and young children are able to track proportional information and use this information to make inferences about the world (Denison, Reed, & Xu, 2013; Duffy, Huttenlocher, & Levine, 2005; McCrink & Wynn, 2007). However, older children show systematic whole number biases when reasoning about discrete proportions, leading to errors such as deciding that 2 out of 3 is less than 4 out of 9 because 2 < 4 even though 2/3 > 4/9 (Hurst & Cordes, 2018). This numerical interference is evident in visual nonsymbolic proportional reasoning prior to formal instruction (Boyer, Levine, & Huttenlocher, 2008; Hurst & Cordes, 2018; Jeong, Levine, & Huttenlocher, 2007) as well as in symbolic fraction learning (Ni & Zhou, 2005). Given the pervasive effects of this numerical bias, it is important to better understand how these effects can be mitigated. In the current study, we investigated whether different types of gesture, a communication and learning tool readily available to teachers and learners alike (Goldin-Meadow, 2015; Novack & Goldin-Meadow, 2017), may affect proportional reasoning by drawing attention either toward or away from discrete whole number and continuous proportional information.

Malleability of proportional reasoning

Some recent work has shown that children's proportional reasoning with nonsymbolic displays is malleable. For example, when children first engage in a continuous proportion task where whole number information is not available (and thus is unable to interfere with proportional reasoning), they perform better on a subsequent discrete proportion task that does have the opportunity for numerical interference (Abreu-Mendoza et al., 2020; Boyer & Levine, 2015; Hurst & Cordes, 2018). That is, practice with proportional reasoning can decrease the numerical interference children experience on subsequent trials. In another study, Hurst and Cordes (2019) showed that the labels used to describe proportional information can also guide children's attention toward the relevant quantity and/or away from the irrelevant information. Using visual displays of proportion, they introduced equivalent proportions (e.g., 3/4, 6/8) using either a single categorical label "blick" or traditional fraction labels of "three fourths" and "six eighths." Based on prior work investigating the use of verbal labels for highlighting relations and category membership (Fulkerson & Waxman, 2007; Fyfe, McNeil, & Rittle-Johnson, 2015; Jamrozik & Gentner, 2020), Hurst and Cordes (2019) hypothesized that the single categorical label would help children attend to proportion and inhibit numerical interference relative to traditional labels. This is exactly what they found in two separate studies with 6-year-old children; seeing equivalent fractions labeled with the same categorical label helped children to inhibit whole number biases and attend to proportional information compared with hearing a traditional fraction label. Despite these findings, it would be impossible, and likely not helpful in the long run, to always use these category-like labels for equivalent fractions because fraction labels are necessary for learning fraction symbols. Thus, another approach to highlighting proportional information and inhibiting numerical interference is necessary.

Gestures as a tool for directing attention

One possible tool for highlighting relevant proportional information, even in the face of irrelevant and interfering discrete numerical information, is the use of hand gestures. Gesture has been shown to support reasoning and teaching in a wide range of domains (Cook, Duffy, & Fenn, 2013; Novack, Goldin-Meadow, & Woodward, 2015; Ping & Goldin-Meadow, 2008; Valenzeno, Alibali, & Klatzky, 2003). One math domain that has been particularly fruitful for studying the effects of gesture is math equivalence. Children make a common error in nonstandard math equivalence problems (e.g., 2 + 3 + 4 = __ + 4) where they add up all the numbers (i.e., putting 13 in the blank) or all the numbers on the left side (i.e., putting 9 in the blank) rather than truly thinking about making the two sides of the equation equal (McNeil, 2007, 2008). This is often described as a procedural heuristic for interpreting the equal sign to mean "add all" rather than a marker of equivalence. However, when children are taught math equivalence problems using gestures that highlight specific strategies, their problem solving improves relative to children who are given the same strategies with speech alone (Broaders, Cook, Mitchell, & Goldin-Meadow, 2007; Congdon et al., 2017; Singer & Goldin-Meadow, 2005). For example, Congdon and colleagues (2017) found that pairing an equalizer strategy in speech (i.e., saying that both sides of the equation must be equal) and an add–subtract strategy in gesture (i.e., pointing to all the numbers on the left and using a "remove" gesture to indicate subtracting the number on the right) was more effective for learning math equivalence than when both strategies were provided in speech.

Furthermore, the effect of gestures may vary based on the particular *kind* of gesture being used. For example, in the context of math equivalence, the benefit of gesture is larger when the information provided in gesture is *different from* the information provided in speech (e.g., describing one possible strategy in speech and simultaneously illustrating a different strategy in gesture) than when both gesture and speech reference the *same* strategy (Singer & Goldin-Meadow, 2005; but see Wakefield & James, 2015). In addition, abstract gestures (e.g., pointing) may facilitate concept generalization more than concrete gestures that mimic the hand movements of actions (e.g., miming physically picking up objects but not actually touching them; Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014), and gestures that are aligned with the relevant concept (e.g., a vertical hand to refer to the middle of a distribution) are more beneficial than gestures that are misaligned with the concept (e.g., a horizontal hand) (Zhang, Givvin, Sipple, Son, & Stigler, 2021). Moreover, these positive effects seem to be unique to gesture in particular and are not seen for equivalent embodied movements that involve action (i.e., moving the hands in a way that acts on the world and moves objects around) (Novack et al., 2014; Wakefield et al., 2019). For example, Novack and colleagues (2014) found that using a grouping gesture (i.e., using a "v" hand shape to point to numbers on the left side of the equation followed by a point to the blank space on the right side of the equation) was more effective than showing the same grouping strategy with actions (i.e., moving physical numbers in a similar way). Together, this work highlights the potentially powerful role for providing information via gesture that can help children to overcome mathematical reasoning errors. In the current study, we aimed to compare two different types of gestures that align with speech but highlight distinct aspects of the stimuli: one that highlights discrete whole number information in the display (*discrete* gesture) and another that highlights the continuous relational information (*continuous* gesture).

Children's own gestures

Not only is gesture powerful as a teaching tool, but also the gestures that individuals themselves produce are important in that they can provide insight into their reasoning (Kelly, Manning, & Rodak, 2008; McNeill, 1992). In particular, in the context of learning, children's gestures have been found to reveal transitional knowledge that is not yet explicit. For example, when learning the verbal counting procedure, preschoolers' spontaneous gestures have been found to reflect their implicit knowledge of the procedure (Gordon, Chernyak, & Cordes, 2019). Moreover, work on math equivalence reveals that when children gesture either spontaneously or as directed, these gestures often reflect different but complementary strategy information from what they say they are using to solve the task, and these inconsistencies may reflect a readiness to learn (Broaders et al., 2007; Perry, Breckinridge Church, & Goldin-Meadow, 1988). In both math equivalence and Piagetian conservation tasks, children who showed a mismatch between their speech and gestures were better able to learn from subsequent instruction than those whose gestures and speech matched (Broaders et al., 2007; Church & Goldin-Meadow, 1986; Perry et al., 1988). In the current study, we took an initial step toward understanding whether children's own gestures in proportional reasoning tasks may reflect their reasoning about the trade-off between numerical and proportional information.

Inhibitory control and math ability

Lastly, in addition to investigating the role of gesture in preventing or encouraging numerical interference, we were also interested in how individual differences in inhibitory control and/or general math ability might play a role in whole number biases found in proportional reasoning tasks. Specifically, these whole number biases have been interpreted as difficulty in inhibiting prepotent numerical responses in the face of continuous proportional information; that is, discrete numerical information *interferes with* the ability to attend to the proportion. Prior work has shown that inhibitory control is correlated with math domains that commonly involve misconceptions (e.g., Abreu-Mendoza et al., 2020; Avgerinou & Tolmie, 2020; Brookman-Byrne, Mareschal, Tolmie, & Dumontheil, 2018; Coulanges et al., 2021; Gómez, Jiménez, Bobadilla, Reyes, & Dartnell, 2015; Kalra, Binzak, Matthews, & Hubbard, 2020; Ren, Lin, & Gunderson, 2019; Vosniadou, 2014). In the case of proportional reasoning specifically, a recent study by Abreu-Mendoza and colleagues (2020) found that children with higher inhibitory control were better able to correctly use proportional information to make judgments about probability even in the presence of competing numerical cues, suggesting that inhibitory control may be an important factor for helping children to overcome their numerical biases. In the current study, we aimed to extend the findings of Abreu-Mendoza et al. (2020) to investigate the robustness of the relation between inhibitory control and proportional reasoning with slightly younger children and with different measures of both proportional reasoning and inhibitory control. In addition, although much of the research on informal proportional reasoning is motivated by an interest in supporting formal mathematical learning, the evidence of a relation between nonsymbolic ratio and formal math is typically found in adults or older children who have already learned formal fractions (e.g., Matthews, Lewis, & Hubbard, 2016; Möhring, Newcombe, & Frick, 2016). No work to date has explored whether nonsymbolic proportional reasoning is related to mathematical processing in children who have yet to receive formal instruction in fractions. Thus, in the current study, we also included a measure of general math knowledge appropriate for these younger children who have not learned formal fractions. This approach allowed us to investigate whether children's informal proportional reasoning may also be related to symbolic and formal math knowledge even prior to learning formal fractions.

The current study

Thus, in the current study, we had three specific research questions. (1) How do different types of gesture influence attention to number in a proportional reasoning task? (2) Do children's own gestures for proportional information relate to their proportional reasoning ability? For both these questions, we predicted that children who produced or saw (via the experimenter) gestures that highlighted continuous proportion would demonstrate lower numerical interference, whereas producing or observing gestures that highlighted discrete parts would be associated with greater numerical interference. (3) Is proportional reasoning related to inhibitory control and math ability? We hypothesized that inhibitory control would be particularly relevant for preventing numerical interference, consistent with other recent work (Abreu-Mendoza et al., 2020), whereas math ability would be related to proportional reasoning more generally (i.e., in the absence of numerical interference), reflecting that performance in this task may reflect the foundations on which later fraction knowledge is built.

Method

Participants

The final sample comprised 135 children aged 5.5–7.5 years (M_{age} = 76 months or 6.30 years; 61 girls and 74 boys). Children were randomly assigned to one of three conditions: Discrete Gesture

(n = 45; M_{age} = 6.29 years; 15 girls and 30 boys), Continuous Gesture (n = 45; M_{age} = 6.34 years; 20 girls and 25 boys), or No Gesture (n = 45; M_{age} = 6.27 years; 26 girls and 19 boys).[3] The sample size was determined a priori based on Hurst and Cordes (2019). Sensitivity analyses using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) suggests that n = 45/condition provides 80% power to detect a relatively small interaction (f = 0.14, which corresponds to approximately d = 0.30) between the between-participant condition and the within-participant trial type, which is the primary analysis. Furthermore, the full sample (N = 135) provides sensitivity to detect correlations as small as .24.

Participants were recruited from the greater Boston, Massachusetts area in the northeastern United States and were tested in a variety of locations, including our campus laboratory (38%–42% of each condition), local museums (2% of each condition), and other off-site locations such as schools and childcare centers (56%–60% of each condition).[4] Parents provided informed written consent, and children provided verbal or written assent, depending on their age. Children received small toys or stickers for participating. All procedures were approved by the Boston College institutional review board.

Demographic data were collected only from the subset of our sample collected in our lab. For these children, parents reported children's race/ethnicity as follows: 76% White, 7% Asian, 0% Black or African American, 0% American Indian or Alaskan Native, 0% Native Hawaiian or Pacific Islander, and 17% mixed race. Approximately 8% reported being Hispanic. Furthermore, of those who reported their education, all mothers had at least a high school diploma, 85% had at least a bachelor's degree, and 57% had a master's or doctorate degree. Although we did not collect demographic information from participants tested outside of the lab, based on the demographics of the schools in which data were collected, we estimate the school-based subset of our sample to be approximately 53% White, 7% Asian, 19% Black or African American, 18% Hispanic, <1% Native American or Alaskan Native, <1% Native Hawaiian or Pacific Islander, and 1% mixed race. We do not expect children tested at local afterschool programs and museums to have substantially different demographics distributions from those tested at schools or in the lab.

Measures and procedure

Participants completed the following tasks in this order: (1) vocabulary assessment, (2) Stroop task (measure of inhibitory control), (3) proportion training (varied by condition), (4) learning verification, (5) match-to-sample task, (6) gesture production task, (7) coloring task, and (8) math assessment. The proportion training, learning verification, match-to-sample task, and math assessment were modeled after the Traditional Label condition of Experiment 2 in Hurst and Cordes (2019).

The primary dependent variables came from the match-to-sample task, which measured children's proportional reasoning ability both with and without the opportunity for numerical interference. The gesture production Task was included as a measure of children's own strategy for representing proportional information via gesture. The coloring task was included to provide another means of investigating children's representation of proportional information visually, but through action, distinct from gesture. The vocabulary assessment, Stroop task, and math assessment were included to ensure there that were no systematic differences across groups on these measures and to investigate how individual differences in proportional reasoning may be uniquely related to inhibitory control (as measured by the Stroop task) and math ability.

Vocabulary assessment

We used the standard protocol for the Woodcock–Johnson III Picture Vocabulary test (Woodcock, McGrew, & Mather, 2001) in which children were shown pictures (up to six on each page) of different

[3] Although not intentional, our conditions differ in the distribution of gender. However, there are no a priori reasons to expect gender differences in proportional reasoning, and comparisons of performance on the current proportional reasoning tasks revealed no significant gender differences in performance in any of the three conditions on Interference trials (ps > .20) and No Interference trials (ps > .08). Thus, we do not include gender in any of our analyses.

[4] Initial analyses comparing children tested at off-site locations with those tested at our campus laboratory showed no significant main or interaction effects on the primary proportional reasoning task across testing locations. Therefore, testing location is no longer considered in the reported analyses.

objects and were asked to name these objects. Children were required to meet a starting criterion of naming six objects correctly in a row and an ending criterion of naming six objects incorrectly in a row. The highest item reached before meeting the ending criterion was used in the analysis. Eight children were excluded from analyses involving vocabulary because it was not administered (n = 3), because it was administered incorrectly (n = 4), or because of parental interference during the task (n = 1).

Stroop task

We used an adapted version of the day–night Stroop (e.g., Gerstadt, Hong, & Diamond, 1994) as a measure of executive function. The task showed good reliability in the current sample with α = .89 calculated with the *psych* package (Revelle, 2020). Children were presented with images and instructed to say the opposite of the image depicted (e.g., say "happy" for a picture of a sad face, say "foot" for a picture of a hand). All participants were taken through an initial introduction and practice followed by 24 test trials. First, children were shown the eight images (happy face, sad face, up arrow, down arrow, hand, foot, open door, and closed door) and were asked which label applied. For example, the experimenter would show the child an image of a foot and ask, "Is it a hand or a foot?" Children were corrected as needed and given feedback to ensure that they knew the labels of all eight images. Then, children were told that they were going to play an opposite game so that when the experimenter showed them a picture, they should say the opposite label. Children were taken through four demonstrations, one for each pair (happy–sad, up–down, hand–foot, and open–closed), where they were shown the image and given the opposite word. Then, children were given 4 practice trials using the remaining cards not used in the demonstration trials during which they were asked to provide the answer, received feedback for their response, and were prompted until they responded correctly. After correctly completing all the practice trials, children were given 24 test trials in a random order (3 of each image) where they were asked to provide the opposite label as quickly as they could.

Children were given 1 point per correct response on the test trials (corrected responses, where children said or started to say the incorrect response but corrected themselves, were not accepted). Due to experimenter error, 4 children received only 22 or 23 trials, so proportion correct out of the number of trials administered is used as the dependent variable. Six children completed the task but were not included in the analyses because their responses could not be scored due to either missing recordings or inaudible responses on the recordings.

Proportion training.

The procedure for proportion training was identical across conditions except for the types of gestures used. Throughout, children sat on the left side of the experimenter so that all gestures could be performed by the experimenter with the right hand. Visual depictions of the condition-specific gestures are provided in Fig. 1, and videos of the gestures are available on our Open Science Framework project page (https://osf.io/43p5g/).

In all conditions, children were first introduced to a character named "Roo" (a toy kangaroo) who "likes shapes with just the right amount of color and just the right amount with no color." The experimenter then brought out an empty gray circle divided into fourths and told children, "I'm going to color three fourths." The experimenter proceeded to color three fourths of the circle, and then emphasized the proportion colored by saying, "See, this is called three fourths. Three [condition-specific gesture highlighting numerator] fourths [condition-specific gesture highlighting denominator]." In the Discrete Gesture condition, the experimenter pointed to each individual unit within the shape. The experimenter would say "three" while pointing three times, once to each of the colored segments in the numerator, and then would say "fourths" while pointing four times, once to each segment in the total shape (i.e., the denominator). In the Continuous Gesture condition, the experimenter would say "three" while dragging the pointer finger along the three colored segments (numerator units) in a single continuous motion and then would say "fourths" while dragging the pointer finger along the entire shape, starting and ending at the top of the circle or going from one end of the rectangle to the other. In the No Gesture condition, the experimenter simply said the label without any gestures. However, to ensure that the two gesture conditions did not take substantially longer than the No Gesture condition (given that the gestures took longer than just saying the label), the experimenter

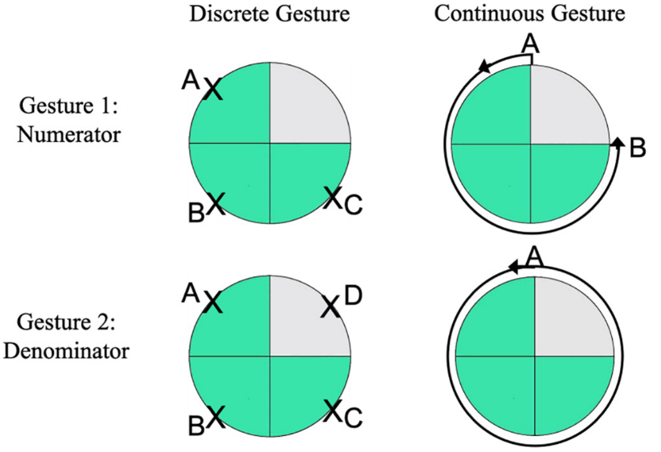

Fig. 1. Visual depiction of the gestures used in the Discrete Gesture (left) and Continuous Gesture (right) conditions to highlight the numerator (Gesture 1; top) and the denominator (Gesture 2; bottom). In the Discrete Gesture condition, the experimenter pointed to each segment of the numerator (A, B, and C; Gesture 1) and then to each segment of the whole shape (A, B, C, and D; Gesture 2). In the Continuous Gesture condition, the experimenter dragged along the green portion (from A to B; Gesture 1) and then dragged along the entire shape (from A to A; Gesture 2).

used the discrete or continuous gesture motions under the table and out of sight of the child to help pace the labeling.

After the shape was introduced with the condition-specific gestures, the shape was then placed on a tray in front of Roo. Next, the experimenter displayed a precolored three-fourths rectangle and introduced it using the same labeling approach and condition-specific gesture as for the three-fourths circle. Finally, the experimenter introduced two shapes (3/8 of a circle and 4/4 of a rectangle) as proportions that Roo would not like and placed them off the tray but still visible (note that no gestures were used for these shapes in any condition). This full procedure was then repeated using shapes with six eighths colored (equivalent fraction to 3/4) as more examples of shapes that Roo likes (with condition-specific gestures) and 6/12 and 2/8 used as counterexamples of shapes that Roo does not like (without gestures). The specific counterexamples were chosen to emphasize that matching based on the numerator alone (3 and 6) or the denominator alone (4 and 8) was not sufficient. To conclude the training, the experimenter reemphasized the amount of color that Roo likes by saying, "Remember that this is the amount of color Roo likes [condition-specific gesture highlighting the colored portion of each circle], and this is also the amount of color Roo likes [condition-specific gesture highlighting colored portion of each rectangle]." In the No Gesture condition, the experimenter just pointed once at each shape Roo liked.

Learning verification

After the training, children completed a verification task to ensure that they learned what Roo liked. The task was presented on a 13-inch Apple laptop via Xojo programming software. On each trial, children were shown two shapes, one on the left and one on the right, with a proportion colored in and were asked which of the two shapes Roo would like. There were a total of 6 trials in two trial types: Learned trials (2 trials), which included the exact values (3/4 and 6/8) and shapes (circles and rectangles) used during training, and Generalized trials (4 trials), all of which were presented as squares (i.e., a novel shape) and included both learned and nonlearned proportions (e.g., 12/16). On each trial, the experimenter recorded children's response by selecting the left or right arrow key (corresponding to the left or right shape).

Match-to-sample task

The match-to-sample task immediately followed learning verification via the same Xojo program. The task was introduced as a new activity without any explicit reminder or indicator to think about what Roo liked or to use a particular proportion strategy. The task followed a match-to-sample procedure in which children were shown a circle with a proportion shaded in green on the upper half of the screen and two rectangles, each with a different proportion shaded in green, presented as options on the lower half of the screen (see Fig. 2). The sample circle became visible 1000 ms before the rectangles became visible. On each trial, children were asked which of the two options matched the sample at the top without any additional instructions about how to best make their match. The experimenter recorded children's response by pressing the left or right arrow key (corresponding to the left or right option). There were 13 trials with two trial types: Interference trials (8 trials) and No Interference trials (5 trials).

On Interference trials, one of the options was an equivalent proportion to the sample but not an exact match (and so did not have the same numerator or denominator), whereas the other option was not an equivalent proportion but matched on a numerical feature. Specifically, the alternative option either had the same numerator and different denominator (Numerator Interference trials; $n = 4$; e.g., sample of 4/10 with 2/5 [proportion match] and 4/6 [numerator match] as options) or had the same denominator and different numerator (Denominator Interference trials; $n = 4$; e.g., sample of 5/7 with 10/14 [proportion match] and 3/7 [denominator match] as options). Given that participants were not explicitly instructed to match on proportion, these trials measured their tendency to focus on proportion instead of number in light of their prior introduction to proportional amounts.

On No Interference trials, one option always matched on proportion and the other option did not match on proportion, but the specific identity varied in two ways. On Exact Match trials, the correct answer was an exact proportional and numerical match (e.g. 3/5 sample with 3/5 [exact match] and 4/12 [foil] options). On Equivalent Match trials, the correct response was an equivalent proportion and so matched on proportion but not number, and the incorrect answer did not match on either proportion or number (e.g., 2/6 sample with 1/3 [proportional match] and 4/7 [foil] options). Again, although there was no explicit instruction to match on proportion, these trials served as a baseline measure of children's tendency to match on proportion when common alternatives (namely, numerical information) were not in competition.

The dependent measure on all trial types was the proportion of trials in which children selected the option that matched the sample on proportion.

Gesture production task

We then assessed children's own use of gesture in order to investigate whether the type of gestures children produced matched the gestures they observed during the demonstration and whether individual differences in children's gesture patterns shed light on their reasoning during the proportional reasoning task. Children were presented with four shapes, one at a time, and were asked to "show me with your fingers how much color Roo would like." Children were shown the following shapes in this order: (1) an empty undivided circle, (2) an empty undivided rectangle, (3) a circle divided into fourths but not colored or filled in, and (4) a rectangle divided into fourths but not colored or filled in. Children's gestures were video-recorded using QuickTime from a laptop pointed toward the table to get a "top view" as well as from a standard tripod video camera. One child did not complete this task and was excluded from these analyses.

Children's videos were coded to identify all gestures used on each trial. Each gesture was coded as being *discrete* or *continuous.* Similar to the gestures used by the experimenters during training, discrete gestures were those that highlighted number such as pointing to individual units within the shape and slicing the shape into a specific number of units. Continuous gestures were those that highlighted multiple segments or sections as one larger unit such as dragging motions and coloring motions. Gestures that were a combination of both discrete and continuous (e.g., slicing a shape into units and then using a coloring motion that went across multiple units simultaneously) or something else entirely (e.g., a slapping motion toward the table or pinching the hands in an upside-down cup motion) were also coded separately. Children were given a score on each trial for the number of gestures of each type and, given our theoretical interest in discrete and continuous gestures specifically,

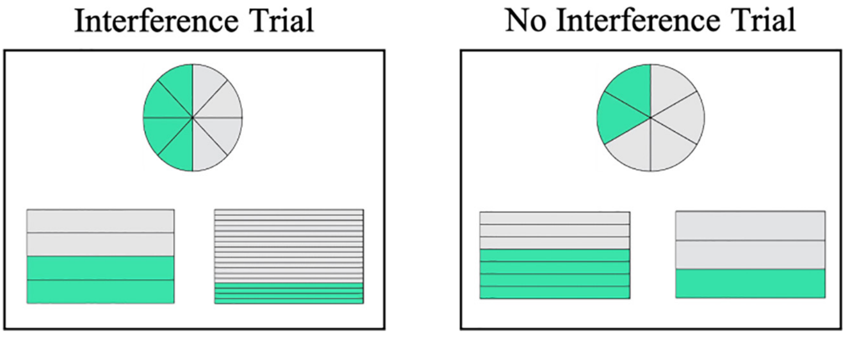

Fig. 2. Example trials from the equivalence match-to-sample task. Left: Interference trial, with the left option matching on proportion (50%) but not on number and the right option matching on number (4 green) but not on proportion (numerator interference). Right: No Interference trial, with the right option matching on an equivalent proportion and the left option not matching.

the proportion of total gestures that were categorized as exclusively discrete and the proportion of total gestures categorized as exclusively continuous are reported and used in the analyses. Two independent researchers coded all gestures, and inter-rater reliability as measured using the *irr* R package (Gamer, Lemon, & Puspendra Singh, 2019) was high for the proportion of gestures that were discrete (intraclass correlation [ICC] = .76) and continuous (ICC = .82). The average across the two raters was used in all analyses because it provides a more accurate measurement (Hallgren, 2012; Shrout & Fleiss, 1979). However, the reported results and conclusions do not change when ratings from a single coder are used instead.

In addition, the two raters who coded gestures (as described above) also transcribed the language used during the execution of each gesture. However, children produced relevant speech with fewer than half of the continuous and discrete gestures (average of 34% and 48% of the continuous and discrete gestures, respectively), likely because the task was not designed to elicit speech given that children were explicitly told to "*show* me with your hands." Thus, given the very low frequency of co-gesture language and our primary interest in gesture specifically, we do not report any additional analyses involving speech and instead focus on children's gesture use.

Coloring task

In addition to the gesture production task, we included a task where children were asked to draw the proportion using a marker to measure their tendency to demonstrate proportion using *action*, in contrast to gesture, and as another measure of their learning of what amount Roo liked. Children were given a marker and were instructed to color the amount that Roo likes on each shape. Divided but uncolored shapes were presented one by one on laminated cardstock that could be washed and reused after each child's use. There were a total of 4 trials presented in the following order: (1) a circle divided into fourths, (2) a rectangle divided into fourths, (3) a circle divided into tenths,[5] and (4) a rectangle divided into twelfths. One child did not complete this task and was excluded from these analyses.

As in the gesture production task, the drawing task was coded with the coloring style for each shape being either entirely *discrete*, entirely *continuous*, or a combination of the two or other. Given the inherent motion of coloring, however, it is worth noting that the code was based on children's treatment of the discrete units in the drawing (i.e., highlighting their discreteness or ignoring them). Discrete coloring styles highlighted separate units such as coloring or outlining each segmented piece within a shape individually. Continuous coloring styles highlighted continuous information such as coloring across the lines that divided the units or outlining multiple units together. Again, two

[5] This trial is an error and should have been divided into twelfths. When divided into tenths, the correct response is 7.5 out of 10 pieces, which children may have been unlikely to realize. Therefore, we kept this trial when coding for strategy but removed it when coding for accuracy.

independent researchers coded each child's drawing with high inter-rater reliability on the proportion of trials that were discrete (ICC = .96) and the proportion of trials that were continuous (ICC = .90), and the average across the two coders was used in all analyses.

In addition to coding coloring strategies, children's responses were recorded based on the number of units children colored in to provide some additional insight into their accuracy. Two raters recorded all trials and disagreed on only 1 trial, which was then coded by a third rater.

Math assessment

Problems were adapted from the Woodcock–Johnson III Applied Problems test (Woodcock et al., 2001). There were 10 problems that covered a range of math domains, including identifying quantities, counting, arithmetic based on images, and arithmetic based on word problems. Children responded verbally, and experimenters live-scored their responses. Children were scored based on the number of problems correct out of 10.

Open data and data analysis

All analyses were done in R 4.0.2 (R Core Team, 2020) with RStudio (RStudio Team, 2016). For data wrangling and organization, we used *dplyr* 1.0.1 (Wickham, Francois, Henry, & Muller, 2018), *tidyr* 1.1.1 (Wickham & Henry, 2018), *readxl* 1.3.1 (Wickham & Bryan, 2019), and *readr* 1.3.1 (Wickham, Hester, & Francois, 2018). Statistical analyses comparing children's behavior between conditions was done using analyses of covariance (ANCOVAs) and *t* tests from the *rstatix* 0.6.0 package (Kassambara, 2020b) and base R. Correlational analyses were also computed using *rstatix* 0.6.0 (Kassambara, 2020b) as well as *ppcor* 1.1 (Kim, 2015) for partial correlations. Data visualizations were created using ggplot2 3.3.2 (Wickham, 2016), ggpubr 0.4.0 (Kassambara, 2020a), and ggbeeswarm (Clarke & Sherrill-Mix, 2017).

All materials, data, and analysis code are provided on the Open Science Framework project page (https://osf.io/43p5g/).

Results

Initial analyses

There were no significant differences across conditions (Continuous Gesture, Discrete Gesture, and No Gesture) in children's vocabulary, inhibitory control, math score, or age (in months): vocabulary, M_{Cont} = 19.4, M_{Disc} = 19.3, $M_{\text{NoGesture}}$ = 19.6, p = .869, partial η^2 = .002; Stroop, M_{Cont} = 0.83, M_{Disc} = 0.86, $M_{\text{NoGesture}}$ = 0.85, p = .776, partial η^2 = .004; math, M_{Cont} = 0.72, M_{Disc} = 0.76, $M_{\text{NoGesture}}$ = 0.76, p = .280, partial η^2 = .019; age, M_{Cont} = 76 months, M_{Disc} = 75 months, $M_{\text{NoGesture}}$ = 75 months, p = .857, partial η^2 = .002. This suggests that the children did not systematically differ on any of these key dimensions across the three conditions. Furthermore, initial exploratory analyses suggest that age did not interact with condition; thus, we include age (in months) as a covariate throughout to account for expected age-related differences in performance (as in Hurst & Cordes, 2019) but do not investigate it further. The pattern of significance throughout is identical when age is not included as a covariate.

Condition differences on proportional reasoning

See Table 1 for means and standard deviations of both proportional reasoning tasks separated by condition.

First, we looked at children's learning of the specific proportional values that Roo liked (i.e., 3/4 and 6/8) using a one-way ANCOVA across the three conditions when controlling for children's age (in months). Overall, children performed very well on the learning verification task and, in line with prior work (Hurst & Cordes, 2019), performance did not significantly differ across conditions, $F(2, 131)$ = 0.08, p = .917, partial η^2 = .001. Unsurprisingly, age was a significant covariate, $F(1, 131)$ = 4.67, p = .032, partial η^2 = .034, with performance increasing as age increased.

Table 1
Means (and standard deviations) on the proportional reasoning tasks.

	Learning verification	Match-to-sample No Interference trials	Match-to-sample Interference trials
No Gesture condition	.80 (.18)	.63 (.20)	.49 (.23)
Continuous Gesture condition	.82 (.17)	.71 (.20)	.51 (.24)
Discrete Gesture condition	.81 (.17)	.68 (.20)	.49 (.24)

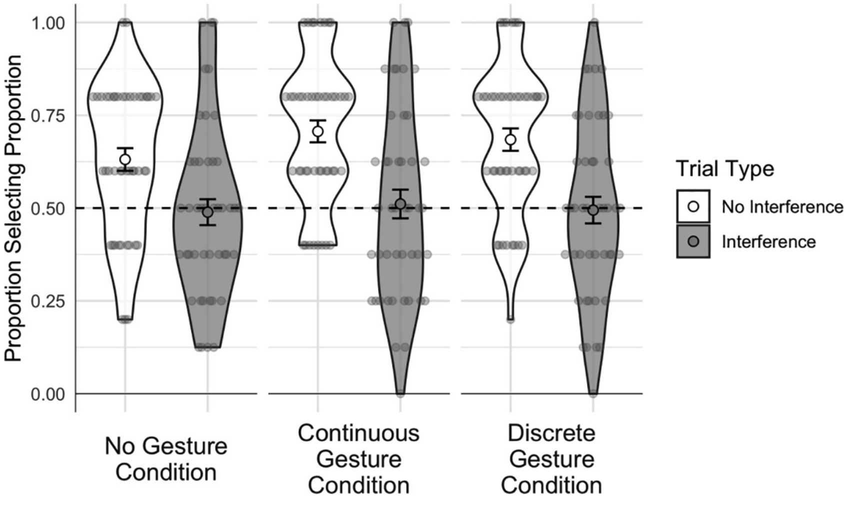

Fig. 3. Proportions correct on Interference trials (left) and No Interference trials (right) in each of the three conditions. White points are means, and error bars are standard errors. Underlying raw data are shown using gray points, with a smoothed kernel density plot showing the distribution of the data.

Second, we examined performance on the Interference and No Interference trials of the match-to-sample task (see Fig. 3) using a 2 (Trial Type: Interference or No Interference) × 3 (Condition: Discrete Gesture, Continuous Gesture, No Gesture) repeated-measures ANCOVA on the proportion of trials where children selected the proportional match, again with age as a continuous covariate. In line with prior research (Hurst & Cordes, 2019), we found a main effect of trial type, $F(1, 131) = 48.39$, $p < .001$, partial $\eta^2 = .270$, with substantially lower proportional matching on the Interference trials ($M = .50$) than on the No Interference trials ($M = .67$). That is, we found evidence of numerical interference in the proportional reasoning task. Importantly, children performed above chance on No Interference trials, $t(134) = 9.98$, $p < .001$, suggesting that they were able to match based on proportion in the absence of competing numerical information. However, they did not perform above chance on Interference trials, $t(134) = -0.09$, $p = .930$. Moreover, analyses did not reveal a significant main effect of condition, $F(2, 131) = 0.84$, $p = .435$, partial $\eta^2 = .013$, or a Trial Type × Condition interaction, $F(2, 131) = 0.42$, $p = .661$, partial $\eta^2 = .006$. In fact, performance on the different trial types was fairly similar across conditions (see Table 1). Thus, we did not find evidence that the experimenter's use of discrete versus continuous gestures during training had a significant impact on children's proportional reasoning. Lastly, again age was a significant covariate, $F(1, 131) = 6.64$, $p = .011$, partial $\eta^2 = .048$, but did not further interact with trial type ($p = .339$).

Table 2
Means (and standard deviations) and ranges of the proportions of gestures within representation types.

	Proportion of gestures that were discrete	Proportion of gestures that were continuous
No Gesture condition	66% (31) [0–100]	25% (27) [0–100]
Continuous Gesture condition	60% (31) [0–100]	32% (28) [0–100]
Discrete Gesture condition	81% (24) [0–100]	15% (20) [0–75]

Note. Ranges are in brackets. Remaining gestures were coded as "combined" or "other."

Children's gesture production

Next, we investigated children's own gesture production to better understand the condition effects (or lack thereof). Overall, as shown in Table 2, children showed substantial individual variation in the type of gestures they produced (although there was also quite a bit of skew in the data). However, on average, most of children's gestures were coded as discrete, followed by continuous. Given that these dependent variables are yoked (i.e., they are proportions of total gestures), we analyzed them separately using one-way ANCOVAs across condition (3: Discrete Gesture, Continuous Gesture, or No Gesture) on the proportion of gestures of that type when controlling for age (however, age was not a significant predictor of gesture production in any of the following ANCOVAs, *p*s > .700, partial η^2s < .001).

When analyzing the proportion of gestures that were discrete, there was a significant difference across conditions, $F(2, 130) = 5.97$, $p = .003$, partial $\eta^2 = .084$, with children using the highest proportion of discrete gestures in the Discrete Gesture condition, followed by the No Gesture condition [vs. Discrete Gesture condition: $t(87) = -2.55$, $p = .012$, $d = 0.54$] and the Continuous Gesture condition [vs. Discrete Gesture condition: $t(87) = -3.48$, $p < .001$, $d = 0.73$]. Discrete gesture use did not differ between the Continuous Gesture and No Gesture conditions, $t(87) = 0.81$, $p = .419$, $d = 0.17$. Similarly, when analyzing the proportion of gestures that were continuous, there was a significant difference across conditions, $F(2, 130) = 5.51$, $p = .005$, partial $\eta^2 = .078$. Children used the highest proportion of continuous gestures in the Continuous Gesture condition, followed by the No Gesture condition, which was not significantly different from the Continuous condition, $t(87) = -1.27$, $p = .208$, $d = 0.27$, and then the Discrete Gesture condition [vs. Continuous Gesture condition: $t(78.4) = 3.43$, $p < .001$, $d = 0.72$]. The Discrete Gesture and No Gesture conditions were also significantly different from each other, $t(78.4) = 2.03$, $p = .046$, $d = 0.43$. Taken together, we see that children's gesture use changed as a function of the gestures they saw during the proportion training, with children who saw discrete gestures producing more discrete gestures and fewer continuous gestures than children in the other two conditions. However, children who saw continuous gestures were not significantly different from those who did not see any gestures, suggesting that observing continuous gestures did not significantly change children's gesture strategies. Moreover, visual inspection of the means makes it clear that children still produced more discrete gestures than continuous gestures across all conditions, suggesting that children's default preference may be toward using discrete gestures.

Children's coloring style

We used the same approach to investigate whether children's coloring of proportional information differed by condition (see Table 3 for descriptive data). Like children's gestures, we saw an overall preference for discrete approaches in children's coloring actions, but unlike children's gestures, we did not find that children's coloring behavior was affected by their condition. There were no significant condition differences in the use of discrete coloring styles, $F(2, 130) = 0.18$, $p = .832$, partial $\eta^2 = .003$, or continuous coloring styles, $F(2, 130) = 0.33$, $p = .719$, partial $\eta^2 = .005$. Age was a significant covariate for both discrete coloring style, $F(1, 130) = 11.22$, $p = .001$, partial $\eta^2 = .079$, and continuous coloring

Table 3
Means (and standard deviations) and ranges of the proportions of action types on the coloring task.

	Proportion of Discrete Coloring trials	Proportion of Continuous Coloring trials
No Gesture condition	57% (34) [0–100]	20% (26) [0–100]
Continuous Gesture condition	52% (34) [0–100]	25% (30) [0–100]
Discrete Gesture condition	56% (38) [0–100]	20% (27) [0–100]

Note. Ranges are in brackets. Remaining coloring actions were coded as "combined" or "other."

style, $F(1, 130) = 9.91$, $p = .002$, partial $\eta^2 = .071$. Specifically, as children's age increased, the proportions of their coloring styles that were discrete increased, whereas the proportions of their coloring styles that were continuous decreased.

The coloring task was coded for how accurately children colored in the proportion of a shape that Roo liked (i.e., an equivalent proportion to 3/4). Nonparametric statistics are reported because only 3 trials were included in these analyses. Accuracy coding on the coloring task revealed that in each of the three conditions, the median accuracy was 2 out of 3, with interquartile range (IQR) = 0. Specifically, 75%, 69%, and 76% of children scored 2/3 in the No Gesture, Discrete Gesture, and Continuous Gesture conditions, respectively. A Kruskal–Wallis rank sum test across the three conditions was not significant, $\chi^2(2) = 3.14$, $p = .208$. Most often, the coloring trial children found to be the most challenging was the rectangle divided into twelfths, with only 7%, 18%, and 4% of children scoring correctly in the No Gesture, Discrete Gesture, and Continuous Gesture conditions, respectively. Thus, children were readily able to produce what Roo liked for denominators they had previously seen, but only a small number of children were able to color the correct proportion on a new denominator value.

Individual differences

Lastly, we explored whether children's proportional reasoning, as well as their numerical interference specifically, was related to individual differences in (a) children's gestures or coloring methods and (b) other skills, specifically their executive function (as measured by the Stroop task) and math ability (see Table 4). Given the lack of condition differences in our primary outcomes, we collapsed across conditions, which also provided increased sensitivity to detect smaller correlations (at 80% power, the smallest reliably detected effect is .24 across all conditions at $N = 134$ and is .40 within condition at $N = 45$ using the *pwr* R package by Champely, 2018). However, for transparency, these correlations separated by condition can be found in the Appendix (Table A1).

Gesture use and coloring patterns

First, we investigated whether individual differences in the use of discrete versus continuous gestures were related to children's proportional reasoning both when numerical interference was possible (Interference trials) and when it was not (No Interference trials). Thus, we performed partial correlations, when controlling for children's age, between performance on the Interference and No Interference trials with the proportion of discrete or continuous gestures (see Table 4 and Fig. 4).[6] Neither children's use of discrete gestures (Fig. 4A) nor their use of continuous gestures (Fig. 4B) was correlated with performance on the No Interference trials. However, we found significant correlations with performance on the Interference trials. Children who produced a higher proportion of discrete gestures were less likely to select the proportional match and thus were more likely to select the numerical match (Fig. 4C), consistent with a whole number bias. Conversely, children who produced a higher proportion of continuous gestures were more likely to select the proportional match and thus were less likely to select the numerical match (Fig. 4D). Thus, the kinds of gestures children produced were asso-

[6] We controlled for age to ensure consistency with the prior analyses. However, the pattern of results is identical with bivariate correlations without controlling for age.

Table 4
Descriptive statistics and correlations among individual difference measures.

			No Interference trials		Interference trials	
	M (*SD*)	*N*	Partial correlation *r*	*p*	Partial correlation *r*	*p*
Gesture production						
Proportion that were discrete	.69 (.30)	134	<.01	.996	−**.24**	**.005**
Proportion that were continuous	.24 (.26)	134	−.01	.894	**.28**	**.001**
Coloring trials						
Proportion that were discretely colored	.55 (.35)	134	.10	.243	.02	.859
Proportion that were continuously colored	.22 (.28)	134	.05	.584	.08	.334
Other measures						
Stroop task	.85 (.19)	129	**.20**	**.024**	−.08	.381
Math assessment	.74 (.14)	135	**.24**	**.005**	.07	.454

Note. All reported *r* coefficients are Pearson's partial correlations when controlling for age (in months). The pattern of correlations is identical with bivariate correlations without controlling for age. Significant correlations are bolded.

ciated with which feature—absolute discrete number or continuous proportional magnitude—they would attend to when both features were available.

However, as seen in Table 4, children's approach to coloring in the proportional shapes was not significantly associated with proportional reasoning on either trial type.

Other measures

Next, we analyzed individual differences in executive function and math ability, when controlling for age (in months) as a proxy for more general differences in cognitive development, to investigate the unique relation between proportional reasoning and these other skills (see Table 4).[7] We found that performance on the math assessment and performance on the Stroop task both were significantly correlated with performance on the No Interference trials when controlling for age. That is, children who were able to match proportion in the absence of competing numerical information also performed better on the math assessment and had higher levels of inhibitory control.

Notably, however, neither math nor inhibitory control scores were correlated with performance on the Interference trials. This is counter to our hypothesis that performance on the Interference trials required inhibiting the prepotent numerical response—exactly what the Stroop task was designed to assess. Thus, although individual differences in math and inhibitory control may be related to proportional reasoning in general, we did not find that they were specifically related to children's attention to number in our task.

Discussion

In the current experiment, we investigated whether different types of gestures may guide children's attention when learning about proportional reasoning and, in particular, whether continuous gestures may be an effective tool for reducing children's whole number biases when engaging in proportional reasoning. Although we did not find that the experimenter's use of gestures during the short training affected numerical interference during subsequent proportional reasoning, we did find that individual differences in children's own gestures were related to their performance on Numerical Interference trials. Taken together, these findings suggest that children's gestures may provide a unique window into their whole number biases when learning about proportion and highlight new questions for better understanding gesture's role in proportional reasoning.

[7] Because age alone might not capture variability in school experience (i.e., this age group is around the time of school entry but varies based on time of year), which may be particularly relevant for executive function and formal math skills, we also confirmed the robustness of these effects when controlling for vocabulary instead, and the pattern was identical.

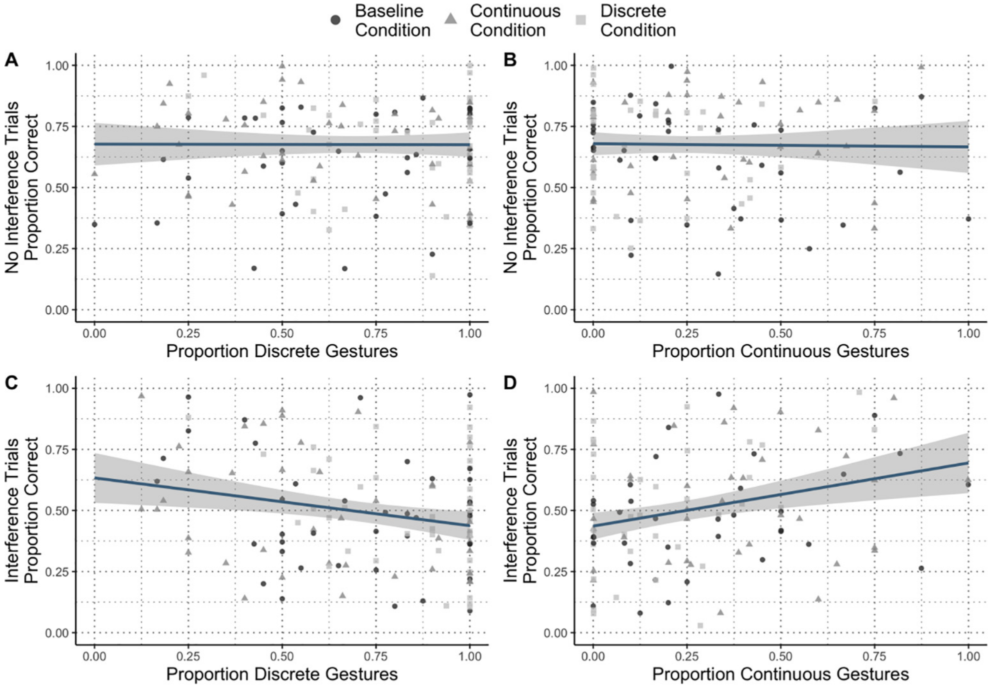

Fig. 4. Bivariate correlation scatterplots of the relations between gesture use (*x* axis) and performance on the equivalence task (*y* axis) with some jitter to see the raw data more easily and condition differentiated by shape and color. Fit lines are linear models with standard errors. (A,B) Nonsignificant relations between performance on the No Interference trials and discrete gesture use (A) or continuous gesture use (B). (C) Negative correlations between the proportions of gestures that are discrete and the tendency to rely on proportions on the Interference trials. (D) Corresponding positive correlations between the proportions of gestures that are continuous and the tendency to rely on proportions on the Interference trials.

The role of gesture in proportional reasoning

Although we found that children's own gestures were informative of their underlying knowledge, our own experimental manipulations did not prove to be effective for influencing their performance on the proportional reasoning task. There are several possible explanations for this pattern that are worth considering.

First, it may be that watching others gesture is not enough to change children's strong biases toward numerical interference, especially through such a brief intervention. This might be particularly relevant given that the experimenter always labeled the proportion using traditional fraction labels, which include number words, and the use of numerical language may have primed children to think in terms of discrete number regardless of whether the gesture was continuous or discrete. In other words, the numerical verbal information may have overshadowed the experimenter's gesture. This might be particularly true given the very brief and experimental nature of the intervention; it may be that with more exposure to both gesture and speech, children would be better equipped to make use of that information. In addition, prior work has highlighted the importance of having children gesture themselves because it can help children to generate new ideas and insights that then make them more likely to benefit from later instruction (Broaders et al., 2007; Kirsh, 2010; Pine, Lufkin, & Messer, 2004; Roth, 2002). However, in the current study, children only gestured *after* performing the proportional reasoning task and we did not directly manipulate children's gestures. That is, even though we did find condition differences in children's gestures, on average children's gestures were still overwhelmingly discrete, and the continuous gesture condition did not have a significant impact on children's gestures. Thus, it could be that having children practice gestures before the proportional reasoning task and encouraging continuous gestures that highlight proportional information may have a stronger impact on their proportional reasoning, thereby lowering the level of numerical interference. Future studies should explore the impact of having children themselves gesture and/or including a more sustained or intensive intervention.

Second, although we chose discrete and continuous gestures that were easy to execute and were aligned with the visual nonsymbolic proportions used in the current study, it may be that the particular gestures used in our study were not the most effective kind of continuous gestures for highlighting proportional information. Future work should investigate what kinds of continuous gestures are readily used to highlight proportional information in more naturalistic and spontaneous contexts than were used here. For example, investigating the kinds of gestures that teachers, parents, and children use to spontaneously communicate about continuous proportional information and then adapting those gestures for direct instruction may lead to more effective take-up of the gestures themselves.

Importantly, however, our results reveal that children's *own* gestures reflect their tendency to rely on numerical versus proportional information in a proportional reasoning match-to-sample task. In line with prior work revealing that children's spontaneous gestures provide insight into their counting and arithmetic abilities (Broaders et al., 2007; Gordon et al., 2019), we find gesture to be an informative tool for better understanding children's errors and signaling when children may need additional support in proportional reasoning. Thus, regardless of the causal direction, it may be informative for educators to pay greater attention to children's gestures during math learning to gain insight into students' approach to the task.

Notably, we did not find a relation between how children colored in a proportional representation and performance on our proportional reasoning task. Although this is in line with other work showing the uniqueness of gesture over action for promoting abstract reasoning and generalization (e.g., Novack et al., 2014; Wakefield et al., 2019), there may be other explanations for the lack of relation with children's coloring approach. For example, all the stimuli used in the coloring task were discrete, which may have biased the way in which children colored. In addition, children are very familiar with coloring and may have defaulted to how they color in any context (i.e., coloring "in the lines") regardless of the proportional information. Lastly, it is worth noting that a reasonable proportion of children's coloring actions were categorized as neither continuous nor discrete (i.e., ~20%–30% were categorized as combined or other), perhaps suggesting a more natural mixing between discrete and continuous coloring, limiting our ability to isolate individual differences in being either discrete or continuous. Regardless of the explanation, however, the unique relation with the types of gestures children produce highlights the need for more work focusing on children's own gestures to better understand this relation and, in particular, whether they might provide a meaningful way to support children's proportional reasoning.

Individual differences in math ability and inhibitory control

Lastly, children's math ability and inhibitory control, as measured by a Stroop task, were significantly related to their proportional reasoning in the absence of numerical interference but not to their numerical interference in particular. Moreover, controlling for age suggests that this pattern is unlikely to be due entirely to age-related changes in cognitive development but instead that there may be unique relations with proportional reasoning and both math ability and inhibitory control.

The relation between math ability and proportional reasoning is particularly notable given that these young children have yet to learn formal fractions and the math task did not include any formal fractions but instead focused on basic counting and arithmetic. This suggests that the relation between nonsymbolic ratio processing and formal math ability found in older children and adults (e.g., Matthews et al., 2016; Möhring et al., 2016) is unlikely to be entirely driven by formal knowledge of fractions. However, the correlational nature of our data precludes causal claims, making it important for future work to investigate the nature of this relation and whether it can be leveraged to support mathematics learning.

The pattern of findings with inhibitory control is in contrast to prior findings with symbolic fractions (Gómez et al., 2015), informal proportional reasoning (Abreu-Mendoza et al., 2020), and other number-based math misconceptions (Ren et al., 2019), which have revealed correlations between inhibitory control and numerical interference in particular. There are a few possibilities for these findings. First, it may be that the use of traditional fraction labels during our task highlighted numerical information more concretely, muting the effect of other individual differences in numerical interference. As noted, prior work suggests that traditional fraction labels, which include references to whole numbers (e.g., "three fourths"), have been found to promote numerical biases in proportional reasoning tasks (relative to categorical labels; Hurst & Cordes, 2019). Second, unlike typical proportional reasoning tasks and those used by Abreu-Mendoza and colleagues (2020), we did not use a cover story that made proportion particularly salient and required attending to proportional information such as a story about probability or juice mixing. That is, although our task was preceded by information about Roo's proportional preferences, the match-to-sample task itself had ambiguous instructions about making the "best match" without an explicit cover story that required the best match to be based on proportion. It may be that inhibitory control is necessary for inhibiting numerical information in contexts where proportion is explicitly necessary but that inhibitory control plays less of a role when the importance of proportion is more ambiguous. That is, it might not be that children needed to inhibit the prepotent response but rather that those who engaged in proportional reasoning simply had a strong preference for proportional information. On the other hand, when it is unambiguously correct to respond with the proportional answer (as in Abreu-Mendoza et al., 2020), children may be more likely to rely on inhibitory control to prevent numerical interference.

Notably, we also found a positive relation between inhibitory control and proportional reasoning in the absence of numerical interference (in contrast to Abreu-Mendoza et al., 2020). It may be that, in light of the ambiguous task instructions in our study, inhibitory control was implicated in our No Interference trials more so than in our Numerical Interference trials. For example, it may be that when a viable alternative response option was not available (in the No Interference trials), attending to proportional information in an ambiguous and numerically salient context required increased executive function (including inhibitory control and working memory) in order to attend to the minimally salient proportional information.

Lastly, it is worth noting that we used a different inhibitory control task than that used by Abreu-Mendoza et al. (2020) that may be tapping slightly different aspects of executive function and inhibition. The hearts and flowers task used by Abreu-Mendoza et al. (2020) requires set shifting in addition to inhibiting a dominant response and activating a subdominant response, whereas the four-pair Stroop task used here does not require set shifting (e.g., Brocki & Tillman, 2014; Montgomery & Koeltzow, 2010). Thus, our distinct pattern of relations may reflect difference in the inhibitory control tasks used and/or differences in the proportional reasoning task. Regardless, the pattern of findings and the various possible interpretations highlight the need for future work investigating proportional reasoning across distinct contexts in order to better understand when children's proportional reasoning relies on inhibitory control and other related skills as well as how these relations may be leveraged to support proportional reasoning.

Conclusion

The current study reveals that children's own gestures provide insight into their proportional reasoning abilities. We did not find that an experimenter's use of gestures during training directly affected children's performance in the proportional reasoning task. However, children's own gesture use was uniquely related to their tendency to rely on numerical versus proportional information in an ambiguous context. This pattern highlights new directions for future research to unpack the role of gesture in proportional reasoning and specifically whether children's own gestures can be leveraged as a powerful tool for facilitating the inhibition of numerical interference in proportional reasoning.

Acknowledgments

We thank the children and families who participated as well as the organizations that helped to facilitate participant recruitment, including the Acton Discovery Museum, the Living Laboratory of the Museum of Science Boston, the Boston Children's Museum, and the schools and childcare programs that partnered with us. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

Materials and data are available on the Open Science Framework (https://osf.io/43p5g/).

Author contributions

Michelle A. Hurst: conceptualization, methodology, software, formal analysis, data curation, writing–original draft, visualization, and supervision; Alyson Wong: formal analysis, investigation, and writing–original draft; Raychel Gordon: conceptualization, methodology, writing–review & editing, supervision, and project administration; Aziza Alam: conceptualization, methodology, investigation, writing–review & editing, and visualization; Sara Cordes: conceptualization, methodology, resources, writing–review & editing, supervision, project administration, and funding acquisition.

Appendix

See Table A1.

Table A1
Descriptive statistics and correlations with gesture use separated by condition.

			No Interference trials		Interference trials	
	M (*SD*)	*N*	Correlation *r*	*p*	Correlation *r*	*p*
Baseline condition						
Proportion of gestures that were discrete	.66 (.31)	44	.21	.164	−.18	.247
Proportion of gestures that were continuous	.25 (.27)	44	−.25	.098	.19	.222
Discrete Gesture condition						
Proportion of gestures that were discrete	.81 (.24)	45	−.039	.798	−.29	.055
Proportion of gestures that were continuous	.15 (.20)	45	.08	.598	**.38**	**.011**
Continuous Gesture condition						
Proportion of gestures that were discrete	.60 (.31)	45	−.19	.218	−.28	.063
Proportion of gestures that were continuous	.32 (.28)	45	.11	.468	**.31**	**.039**

Note. All reported *r* coefficients are bivariate Pearson's *r*. Significant correlations are bolded.

References

Abreu-Mendoza, R. A., Coulanges, L., Ali, K., Powell, A. B., & Rosenberg-Lee, M. (2020). Children's discrete proportional reasoning is related to inhibitory control and enhanced by priming continuous representations. *Journal of Experimental Child Psychology, 199*, 104931. https://doi.org/10.1016/j.jecp.2020.104931.

Avgerinou, V. A., & Tolmie, A. (2020). Inhibition and cognitive load in fractions and decimals. *British Journal of Educational Psychology, 90*(S1), 240–256.

Boyer, T. W., & Levine, S. C. (2015). Prompting children to reason proportionally: Processing discrete units as continuous amounts. *Developmental Psychology, 51*(5), 615–620.

Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: Where young children go wrong. *Developmental Psychology, 44*(5), 1478–1490.

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General, 136*(4), 539–550.

Brocki, K. C., & Tillman, C. (2014). Mental set shifting in childhood: The role of working memory and inhibitory control: The role of WM and inhibition in childhood set shifting. *Infant and Child Development, 23*(6), 588–604.

Brookman-Byrne, A., Mareschal, D., Tolmie, A. K., Dumontheil, I., & De Smedt, B. (2018). Inhibitory control and counterintuitive science and maths reasoning in adolescence. *PLoS One, 13*(6), e0198973. https://doi.org/10.1371/journal.pone.0198973.

Champely, S. (2018). pwr: Basic functions for power analysis (R package Version 1.2-2) [computer software]. https://CRAN.R-project.org/package=pwr.

Breckinridge Church, R., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23*(1), 43–71.

Clarke, E., & Sherrill-Mix, S. (2017). ggbeeswarm: Categorical scatter (violin point) plots (R package, Version 0.6.0) [computer software]. https://CRAN.R-project.org/package=ggbeeswarm.

Congdon, E. L., Novack, M. A., Brooks, N., Hemani-Lopez, N., O'Keefe, L., & Goldin-Meadow, S. (2017). Better together: Simultaneous presentation of speech and gesture in math instruction supports generalization and retention. *Learning and Instruction, 50*, 65–74.

Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development, 84*(6), 1863–1871.

Coulanges, L., Abreu-Mendoza, R. A., Varma, S., Uncapher, M. R., Gazzaley, A., Anguera, J., & Rosenberg-Lee, M. (2021). Linking inhibitory control to math achievement via comparison of conflicting decimal numbers. *Cognition, 214*, 104767. https://doi.org/10.1016/j.cognition.2021.104767.

Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-olds. *Developmental Psychology, 49*(2), 243–249.

Duffy, S., Huttenlocher, J., & Levine, S. (2005). It is all relative: How young children encode extent. *Journal of Cognition and Development, 6*(1), 51–63.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition, 105*(1), 218–228.

Fyfe, E. R., McNeil, N. M., & Rittle-Johnson, B. (2015). Easy as ABCABC: Abstract language facilitates performance on a concrete patterning task. *Child Development, 86*, 927–935.

Gamer, M., Lemon, J., & Puspendra Singh, I. F. (2019). irr: Various coefficients of interrater reliability and agreement (R package Version 0.84.1) [computer software]. https://CRAN.R-project.org/package=irr.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3½–7 years old on a Stroop-like day–night test. *Cognition, 53*, 129–153.

Goldin-Meadow, S. (2015). From action to abstraction: Gesture as a mechanism of change. *Developmental Review, 38*, 167–184.

Gómez, D. M., Jiménez, A., Bobadilla, R., Reyes, C., & Dartnell, P. (2015). The effect of inhibitory control on general mathematics achievement and fraction comparison in middle school children. *ZDM, 47*(5), 801–811.

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.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*(1), 23–34.

Hurst, M. A., & Cordes, S. (2018). Attending to relations: Proportional reasoning in 3- to 6-year-old children. *Developmental Psychology, 54*(3), 428–439.

Hurst, M. A., & Cordes, S. (2019). Talking about proportion: Fraction labels impact numerical interference in non-symbolic proportional reasoning. *Developmental Science, 22*, e12790.

Jamrozik, A., & Gentner, D. (2020). Relational labeling unlocks inert knowledge. *Cognition, 196*, 104146. https://doi.org/10.1016/j.cognition.2019.104146.

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*(2), 237–256.

Kalra, P. B., Binzak, J. V., Matthews, P. G., & Hubbard, E. M. (2020). Symbolic fractions elicit an analog magnitude representation in school-age children. *Journal of Experimental Child Psychology, 195*, 104844. https://doi.org/10.1016/j.jecp.2020.104844.

Kassambara, A. (2020a). ggpubr: "ggplot2" based publication ready plots (Version 0.2.5) [computer software]. https://CRAN.R-project.org/package=ggpubr.

Kassambara, A. (2020b). rstatix: Pipe-friendly framework for basic statistical tests (Version 0.4.0) [computer software]. https://CRAN.R-project.org/package=rstatix.

Kelly, S. D., Manning, S. M., & Rodak, S. (2008). Gesture gives a hand to language and learning: Perspectives from cognitive neuroscience, developmental psychology and education. *Language and Linguistics Compass, 2*, 569–588.

Kim, S. (2015). ppcor: Partial and semi-partial (part) correlation (R package, Version 1.1) [computer software]. https://CRAN.R-project.org/package=ppcor.

Kirsh, D. (2010). Thinking with external representations. *AI & Society, 25*(4), 441–454.

Matthews, P. G., Lewis, M. R., & Hubbard, E. M. (2016). Individual differences in nonsymbolic ratio processing predict symbolic math performance. *Psychological Science, 27*(2), 191–202.

McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science, 18*(8), 740–745.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

McNeil, N. M. (2007). U-shaped development in math: 7-year-olds outperform 9-year-olds on equivalence problems. *Developmental Psychology, 43*(3), 687–695.

McNeil, N. M. (2008). Limitations to teaching children 2 + 2 = 4: Typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development, 79*(5), 1524–1537.

Möhring, W., Newcombe, N. S., & Frick, A. (2016). Using mental transformation strategies for spatial scaling: Evidence from a discrimination task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 1473–1479.

Montgomery, D. E., & Koeltzow, T. E. (2010). A review of the day–night task: The Stroop paradigm and interference control in young children. *Developmental Review, 30*(3), 308–330.

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.

Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: Using the hands to learn math. *Psychological Science, 25*(4), 903–910.

Novack, M. A., & Goldin-Meadow, S. (2017). Gesture as representational action: A paper about function. *Psychonomic Bulletin & Review, 24*(3), 652–665.

Novack, M. A., Goldin-Meadow, S., & Woodward, A. L. (2015). Learning from gesture: How early does it happen? *Cognition, 142*, 138–147.

Perry, M., Breckinridge Church, R., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*(4), 359–400.
Pine, K. J., Lufkin, N., & Messer, D. (2004). More gestures than answers: Children learning about balance. *Developmental Psychology, 40*(6), 1059–1067.
Ping, R. M., & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology, 44*(5), 1277–1287.
R Core Team (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Ren, K., Lin, Y., & Gunderson, E. A. (2019). The role of inhibitory control in strategy change: The case of linear measurement. *Developmental Psychology, 55*(7), 1389–1399.
Revelle, W. (2020). psych: Procedures for personality and psychological research (Version 2.0.9) [computer software]. https://CRAN.R-project.org/package=psych.
Roth, W.-M. (2002). From action to discourse: The bridging function of gestures. *Cognitive Systems Research, 3*(3), 535–554.
RStudio Team (2016). *RStudio: Integrated Development for R*. Boston: RStudio Inc..
Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428.
Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science, 16*(2), 85–89.
Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, 28*(2), 187–204.
Vosniadou, S. (2014). Examining cognitive development from a conceptual change point of view: The framework theory approach. *European Journal of Developmental Psychology, 11*(6), 645–661.
Wakefield, E. M., Foley, A. E., Ping, R., Villarreal, J. N., Goldin-Meadow, S., & Levine, S. C. (2019). Breaking down gesture and action in mental rotation: Understanding the components of movement that promote learning. *Developmental Psychology, 55*(5), 981–993.
Wakefield, E. M., & James, K. H. (2015). Effects of learning with gesture on children's understanding of a new language concept. *Developmental Psychology, 51*(8), 1105–1114.
Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.
Wickham, H., & Bryan, J. (2019). readxl: Read Excel files (Version 1.3.1) [computer software]. https://CRAN.R-project.org/package=readxl.
Wickham, H., Francois, R., Henry, L., & Muller, K. (2018). dplyr: A grammar of data manipulation (Version 0.7.7) [computer software]. https://CRAN.R-project.org/package=dplyr.
Wickham, H., & Henry, L. (2018). tidyr: Easily tidy data with "spread" and "gather" functions (Version 0.8.1) [computer software]. https://CRAN.R-project.org/package=tidyr.
Wickham, H., Hester, J., & Francois, R. (2018). readr: Read rectangular text data (Version 1.3.1) [computer software]. https://CRAN.R-project.org/package=readr.
Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Itasca, IL: Riverside.
Zhang, I. (., Givvin, K. B., Sipple, J. M., Son, J. Y., & Stigler, J. W. (2021). Instructed hand movements affect students' learning of an abstract concept from video. *Cognitive Science, 45*(2). https://doi.org/10.1111/cogs.v45.210.1111/cogs.12940.