Journal of Experimental Child Psychology 192 (2020) 104774

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

How many apples make a quarter? The challenge of discrete proportional formats

Kreshnik N. Begolli [a,*], Julie L. Booth [b], Corinne A. Holmes [c], Nora S. Newcombe [b]

[a] California State University, Long Beach, Long Beach, CA 90840, USA
[b] Temple University, Philadelphia, PA 19122, USA
[c] Trinity College Dublin, Dublin 2, Ireland

ARTICLE INFO

Article history:
Received 10 May 2019
Revised 13 August 2019
Available online 2 January 2020

Keywords:
Proportional reasoning
Spatial cognition
Mathematical development
Rational numbers
Individual differences
Nonsymbolic reasoning

ABSTRACT

Proportional judgments are easier for children in continuous formats rather than discretized ones (e.g., liquid in a beaker vs. in a beaker with unit markings). Continuous formats tap a basic sense of approximation magnitude, whereas discretized formats evoke erroneous counting strategies. On this account, truly discrete formats with separated objects should be even harder. This study (N = 565 7- to 12-year-old children) investigated that prediction. It also examined whether the format effects vary with children's fraction knowledge (FK; part–whole relations, computation, and fraction number line estimation). As found previously, discretized formats were more challenging than continuous ones; as predicted, discrete formats were yet harder. The format effect interacted with FK. Low-FK children were above chance only with continuous formats, medium-FK children struggled with discrete formats only, and high-FK children did well with all three formats.

Introduction

Proportional reasoning is key for mathematical and scientific achievement as well as for navigating everyday life situations (National Council of Teachers of Mathematics, 2000; National Mathematics Advisory Panel, 2008; Siegler, Fazio, Bailey, & Zhou, 2013; Spinillo & Bryant, 1991). At its core,

* Corresponding author.
E-mail address: keko.begolli@csulb.edu (K.N. Begolli).

https://doi.org/10.1016/j.jecp.2019.104774
0022-0965/

proportional reasoning is the ability to reason about part–whole relations—both symbolic and nonsymbolic (Boyer, Levine, & Huttenlocher, 2008; Fischbein & Gazit, 1984; Howe, Nunes, & Bryant, 2011; Piaget & Inhelder, 1974). Proportional reasoning is common in professional contexts such as medicine (e.g., assessing antecedents of disease), anthropology (e.g., comparing bone proportions of human fossils with modern *Homo sapiens*), physics (e.g., density, temperature, velocity), chemistry (e.g., chemical compositions), and economics (e.g., currency exchange). It is also used in everyday contexts such as shopping (e.g., buy two and get one free), gas usage (e.g., miles per gallon), and cooking (e.g., three eggs for two portions).

We now have a good deal of knowledge concerning how children and adults represent and manipulate proportions (DeWolf, Bassok, & Holyoak, 2015; Fischbein & Gazit, 1984; Hoffrage, Gigerenzer, Krauss, & Martignon, 2002; Howe et al., 2011; Hurst & Cordes, 2018; Mazzocco & Devlin, 2008; Piaget & Inhelder, 1975; Siegler, Thompson, & Schneider, 2011; Spinillo & Bryant, 1999; Tian & Siegler, 2018; Van Dooren, Lehtinen, & Verschaffel, 2015; Wason, 1968; Zhu & Gigerenzer, 2006). There is consensus that children can reason about proportions earlier in development if they are presented nonsymbolically (i.e., spatially/visually) rather than symbolically (i.e., numerically). Most studies with children have used a proportional equivalence task in which children match equivalent proportions going from a smaller to a larger absolute size or vice versa (e.g., matching spatialized ratios of 1:2 [1 part chocolate syrup and 2 parts milk] to 2:4 [2 parts chocolate syrup and 4 parts milk]) (e.g., Boyer & Levine, 2012, 2015).

Children are more successful when presented with proportions in continuous formats (e.g., liquid in a beaker) than in discretized formats (e.g., liquid in a beaker with markings) (Boyer & Levine, 2012, 2015; Jeong, Levine, & Huttenlocher, 2007; Spinillo & Bryant, 1999). Visual demarcations in the discretized formats have been argued to prompt children to engage in erroneous counting strategies. For example, they may match a ratio of 3:4 (juice to water) with 3:6 because of the shared numerator (absolute or extensive amount of juice), or they may match 3:9–4:8 because of the shared extensive (absolute) size (3 + 9 = 12 and 4 + 8 = 12) of the "container" (Boyer & Levine, 2012, 2015; Boyer et al., 2008). Counting seems to disengage children from a more intuitive manner of representing proportions as analogue magnitudes (e.g., relations between spatial areas) (Barth, Baron, Spelke, & Carey, 2009; Clearfield & Mix, 1999; Leibovich, Katzin, Harel, & Henik, 2017). Continuous formats allow them to maintain focus on intensive (relative or proportional) magnitude representations (Jirout & Newcombe, 2019), which are arguably more automatic and basic than integers (Leibovich et al., 2017). Counting is also likely to impede children from engaging in the advantageous heuristic of spatial scaling—the ability to mentally enlarge or shrink spatial representations (e.g., mentally enlarging a pocket map to navigate a real-world street).

Spatial scaling appears to be broadly involved in thinking about magnitude concepts and proportional reasoning (Boyer & Levine, 2012; Möhring, Frick, & Newcombe, 2018; Möhring, Newcombe, & Frick, 2014, 2015; Möhring, Newcombe, Levine, & Frick, 2016; Vasilyeva & Huttenlocher, 2004; Vasilyeva, Duffy, & Huttenlocher, 2007). Spatial scaling and proportional scaling recruit overlapping cognitive processes. Scaling requires understanding that the distances on a pocket map and in a real-word space maintain a constant proportion (Möhring et al., 2018). Evidence from both map scaling (e.g., scaling relative positions on a map between the map and a target) and proportional scaling (e.g., scaling between two equivalent proportions) suggests that errors increase with increases in the magnitude of the scale factor (Boyer & Levine, 2012; Vasilyeva & Huttenlocher, 2004). These errors seem to occur because mentally transforming one proportion to match another is a cognitively taxing process. Greater transformations in other spatial tasks—such as angular degree in mental rotation (Shepard & Metzler, 1971) and path length in navigation (Kosslyn, 2005)—have consistently been shown to impose higher cognitive costs, in terms of error rates and/or response latencies, due to their demands on dynamic mental imagery. In addition, children's (and adults') bias toward extensive (absolute) mental representation of magnitudes seems to increase with scaling (or higher quantities) and further deter them from maintaining (and/or encoding) the intensive quantity of the target proportion, exacerbating their errors (Jäger & Wilkening, 2001; for review, see Newcombe, Möhring, & Frick, 2018).

On these accounts, entirely discrete representations, in which every unit representing the parts is separate and spatially separated from the whole, may be very challenging. Although discretized representations seem to lead children to count, the Gestalt properties of entirely separated units that each have a "common fate" should be much more likely to evoke counting (Mandler & Shebo, 1982; Mix, Huttenlocher, & Levine, 2002). Thus, children may find it difficult to represent discrete formats as proportions (Singer-Freeman & Goswami, 2001) and may attempt to count (due to the representations being discrete). Even if children attempt to invoke a proportional (intensive) representation, if they simultaneously count, they will be distracted.

Although prior studies have sometimes called discretized formats "discrete" (Boyer & Levine, 2012), offering a choice between a focus on continuous magnitude and counting demarcated units is different from inviting the counting of spatially separated objects. Children may be more tempted to count such objects, which are the focus of a great many counting exercises, instead of computing magnitude ratios over area or volume, and they may be less likely to use dynamic mental imagery to approximate relative magnitude (Barth et al., 2009; Clearfield & Mix, 1999). Understanding the distinction between discretized and discrete displays could provide important insight into whether the Gestalt of discrete perceptual displays that "force" children to count leads them to make more proportional reasoning errors than displays that can more easily be perceived to represent analogue magnitudes.

Adult data provide further insight into discrete formats. In a series of experiments with adults, DeWolf et al. (2015) used all three representations (continuous, discretized, and discrete), although not on proportional equivalence tasks. They asked about format preferences when matching to a decimal (e.g., .75) or fraction (e.g., 3/4) (Experiment 1) or when evaluating ratio relationships or solving relational tasks (Experiments 2–4). They found that adults prefer discrete and discretized formats for representing fractions, for evaluating ratio relationships with fractions and decimals, and for solving problems involving fractions (DeWolf et al., 2015). Of course, such preferences might not shed light on how children gain adult understanding.

In sum, previous studies have not systematically compared children's performance with continuous, discretized, and discrete formats on a proportional reasoning task (see Fig. 1). Conducting such comparisons was the first purpose of the current study. Investigating the three formats together in samples of children also allows a systematic examination of interactions with other task features. For example, Boyer and Levine (2012) found that increases in scale factor led to decreases in performance, perhaps because larger proportional transformations imposed greater cognitive demands. Yet the relation between format and scale factor might not be consistent across all three formats. In addition, previous work suggests that children are more accurate when the foil constitutes a denominator match with the target proportion than when the foil matches its numerator. In the current study, we systematically compared the three formats at three scaling magnitudes (1.5/.67, 2/.50, and 3/.33) and controlled for directionality (enlarging or shrinking) and foil type (numerator or denominator).

A second purpose of this study involved the relation of format effects to formal knowledge of symbolic fractions—that is, part–whole knowledge, symbolic manipulations, and fraction number line estimation. Prior studies have investigated the developmental trajectories of proportional reasoning and related spatial skills by grouping children based on age or grade level (Boyer & Levine, 2012; Piaget & Inhelder, 1975; Spinillo & Bryant, 1999). However, age and grade level might not always serve as optimal proxies for developmental paths in diverse populations because environmental influences (e.g., socioeconomic status) may promote or slow cognitive and academic growth (Fryer & Levitt, 2004; Hanushek & Woessmann, 2012; Levine, Vasilyeva, Lourenco, Newcombe, & Huttenlocher, 2005). Training can greatly influence performance of spatial representations of proportions (Goswami, 1995; Singer-Freeman & Goswami, 2001). Thus, grouping children by age or grade may combine children at different developmental levels into the same category, making it difficult to detect progressions in proportional ability. An alternative means of studying developmental trajectories is to group children by levels of the target skill under study instead of age or grade level. We used fraction knowledge, which has been found to be closely related to proportional equivalence judgments (Hansen et al., 2015). We grouped children into tertiles of low, medium, and high performers based on their prior knowledge of number line estimations and part–whole concepts and procedures. Using tertiles avoids the problems with median splits, in which children just below or above the median may be quite comparable, while maintaining statistical power to detect interactions.

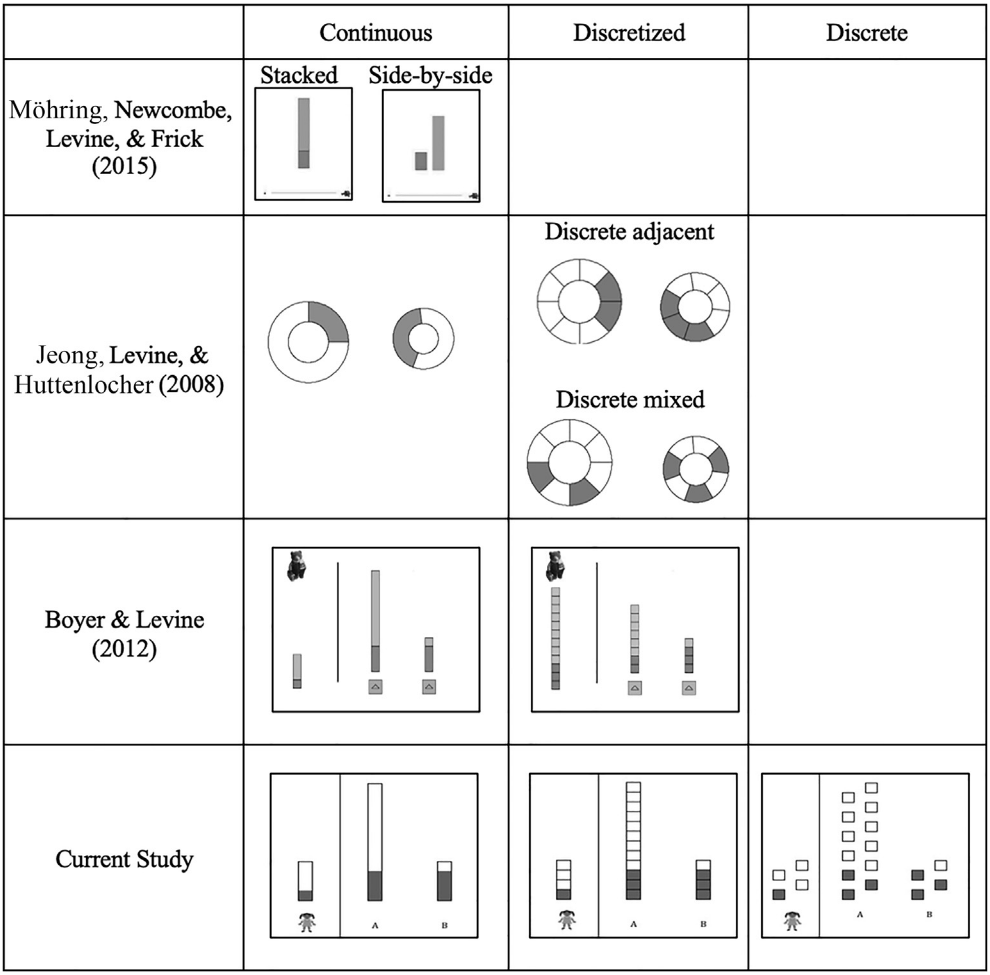


Fig. 1. A contrast of stimuli previously used in the literature and our materials, all of which examine children's performance on proportional equivalence tasks with visually depicted proportions.

Method

Participants

Participants were 565 children (268 boys and 293 girls; 4 children did not report their gender) recruited from seven schools (five private [343 children] and two public charter [222 children]) in the metropolitan area of Philadelphia, Pennsylvania, in the eastern United States. The data were collected in May and June of 2017. There were 182 participants in third grade (92 boys and 89 girls, 1 missing gender information), 194 in fourth grade (98 boys and 94 girls, 2 missing gender information), and 189 in fifth grade (78 boys and 110 girls, 1 missing gender information). The mean age at each grade level was M_3 = 8;6 (years;months), M_4 = 9;7, and M_5 = 10;7. The schools from which the children were drawn ranged from low-income inner-city schools (90% African American families) to those with children from working-class, middle-class, and upper-class families (80% White families). Based on demographic data from school websites, we estimated the overall sample as approximately 62% White, 28% Black, 5% Hispanic, 3% Asian, and 2% multiracial or other. Approximately 44% of children were eligible for free or reduced-cost lunch. Principals of at least 30 private and charter schools were asked via e-mail to participate in the study, and all children in those schools were eligible to participate. All children had parental consent and gave individual assent to participate. Participants who were missing more than half of their data for each measure were excluded from analyses, but such missing information was rare: 3 participants in the proportional equivalence task and 1 participant in the fractions concepts measure (4 participants in total).

Procedure

Participants were administered the paper-and-pencil-based proportional equivalence task and mathematics assessment during their regular class hours in their individual classrooms, in neighboring classrooms, or in the school cafeteria. First, children in all the format types read the instructions together with the experimenter, who introduced them to Goldilocks through the first page of the packet (see Appendix A). The experimenter explained that Goldilocks likes her chocolate milk "just right," and the container(s) above Goldilocks showed how Goldilocks likes her chocolate milk to taste. The experimenter guided children to follow the instruction page (see Appendix A) to see which parts represented milk and which parts represented chocolate syrup; no gestures were used and, as can be seen from Appendix A, the arrows do not point at any individual unit in the discrete format type. Children were instructed to help Goldilocks choose the correct chocolate syrup/milk mixture that would taste "just right" and selected one of two options. To ensure that children understood they should not be matching based on one of the common misconceptions stemming from counting strategies and that they understood this was a proportional task, in addition to language that was previously used, we also added an example with a correct answer. After the experimenter gave an example of the task (see Appendix A), children completed 24 proportional equivalence items, 6 per page, and then continued on to the mathematics assessment. When they reached the end of the packet, the last page asked children to write down their age, gender, and grade level. If children asked questions about the proportional equivalence task or the mathematics problems, the experimenter read the instructions or problem prompt together with the children.

Design

Proportional equivalence task.

Within each classroom and grade level, children were randomly assigned to one of three proportional equivalence tasks that contained continuous, discretized, or discrete spatial representations (see Fig. 1). Trial types were identical across spatial representations and were selected from Boyer and Levine (2012, Experiment 1) (see Appendix B). There were 24 randomly sequenced trials composed of six trial types that varied by scale factor and scaling direction. The scale factor increased when scaling up or decreased when scaling down by three separate magnitudes. When scaling up the magnitude of the proportion increased by 1.5, 2, and 3, and when scaling down the magnitude decreased by .33, .50, and .67. Each magnitude factor was presented four times, with half of the trials scaled up (e.g., 1/3–3/9) and half scaled down (e.g., 3/9–1/3). Scaling up and scaling down were at a symmetrical magnitude (e.g., .33 vs. 3 both are at 3:1). As in Boyer and Levine (2012), all items involved fractions that reduced to 1/3, 1/4, 2/3, or 3/4 chocolate syrup and milk parts.

The foil alternative was always on the opposite side of the half-mark boundary as the target proportion to facilitate comparison between the two choice alternatives (Spinillo & Bryant, 1991). Foils were manipulated such that half of the trials involved a numerator match with the target and half involved a denominator match (see Appendix B). Numerator and denominator foils were evenly distributed such that within the 4 trials per scale factor, 2 contained numerator foils and 2 contained denominator foils. The denominator and numerator foils were chosen because previous work has shown that children tend to match based on absolute size of the container or number of units (i.e., chocolate syrup + milk, denominator match) or the absolute size or number of units of the factor believed to change the taste (i.e., chocolate syrup, numerator match) rather than the proportional relation (Boyer & Levine, 2012, 2015; Boyer et al., 2008; Piaget & Inhelder, 1975).

Fraction assessment

All participating children at all grade levels completed the same mathematics problems. There were 61 items in total, with 45 items measuring formal fraction knowledge concepts and 16 items measuring fraction number line estimation (see Appendix C for examples). We standardized all items and combined the standardized scores from both part–whole and number line estimation tasks into a composite score of overall fraction knowledge. The Cronbach's alpha reliability coefficient of all items combined was .921 across all grades (.901 at third grade, .920 at fourth grade, and .929 at fifth grade).

Fraction concepts. In total, 11 items measured children's understanding of part–whole representations (Hecht & Vagi, 2010, 2012), namely to shade an area based on a fraction (3 items: 1/3, 2/8, and 1/2), represent a shaded area with symbolic fraction notation (3 items: 3/5, 2/4, and 1/4), and circle "yes" or "no" to fraction addition equations involving the sum of two discretized picture representations of fractions (5 items; e.g., 1/3 + 2/6 = 4/6).

In total, 20 items measured children's understanding of fraction equivalence or their ability to compare fraction magnitudes (Schumacher, Namkung, Malone, & Fuchs, 2012; Siegler et al., 2011), where children needed to circle which fractions are equal to 2/4 by circling "equal" or "not equal" (6 items: 1/2, 2/3, 3/6, 3/4, 4/4, and 4/8), circle which fractions are "larger" or "smaller" than 4/12 (4 items: 4/10, 4/16, 2/12, and 8/12), and mark greater than (>), less than (<), or equal to (=) on two fractions (10 items: 2 items with equal numerators, 2 items with equal denominators, 2 items with equal fractions, 2 items with number symbol magnitudes that matched the size of the proportion such that greater numbers in the fraction corresponded to a greater proportion and vice versa, and 2 items where there was a mismatch between the number magnitude and proportion).

In total, 13 items were nested within six word problems and were designed to assess children's conceptual understanding of fractions (Begolli & Richland, 2016; Brown & Quinn, 2006; Paik & Mix, 2003) involving fraction subtraction (6 items), division (1 item), comparison (5 items), and part–whole picture → symbol representation (1 item).

There was 1 item that measured children's ability to subtract mixed numbers (2–4/5–1/5 = ?).

Fraction concept items were scored for accuracy; if correct, children received 1 point per item and the sum score of all items was used in further analyses. Alpha reliability was .85 overall (.87 at third grade, .88 at fourth grade, and .90 at fifth grade).

Fraction number line estimation. For fraction number line (FNL) estimation tasks (drawn from Fuchs et al., 2013; Siegler et al., 2011), children were first shown how to indicate a fraction's location on a number line for two fraction symbols (1/2 and 5/6) and one fraction word (half) and were then asked to solve 16 items on their own; of these, 10 items involved fraction symbols (1/2, 3/4, 1/4, 2/4, 1/3, 2/5, 3/8, 2/3, 9/10, and 1/7) and 6 items involved fraction words (half, a quarter, two quarters, a third, three quarters, and an eighth), which were displayed at the halfway point on top of the number line. The number line was labeled at each end point with 0 and 1 and was 85 mm long. Accuracy on the number line was assessed according to previous work using the percentage absolute error (PAE) (Booth & Siegler, 2006; Opfer & Siegler, 2007; Siegler et al., 2011), where PAE = (child's response − correct response)/(size of number line). The reliability was .84 overall (.82 at third grade, .84 at fourth grade, and .86 at fifth grade).

Analysis plan

We used linear mixed models using STATA 13 (StataCorp, College Station, TX, USA). Because individual children in the study were nested in classrooms and schools, we first computed intraclass correlations (ICCs) on the dependent variable for Level 2 (classrooms; ICC = .029, p = .19) and Level 3 (schools; ICC = .084, p < .01). The ICCs suggest that it was not necessary to consider nesting by classroom; however, nesting by school should be accounted for in our model. Because the number of schools at Level 3 (n = 7) was not sufficient to support a nested model, we instead placed a variable indicating school for each student as a control factor in the model. The final model was a crossed-random-effects model in the form $Y_i = X_iB + Z_ib_i + e_i$, where Y_i represents the values of children's proportional reasoning task for each i^{th} child, X_i denotes the independent variables (between participants:

format, fraction knowledge, and *grade*; within participants: *foil type, direction,* and *magnitude*; control: *gender*) and *B* denotes their respective fixed-effects beta weight estimates, Z_i denotes the random effects (*school* and *child*) and b_i denotes their estimates, and e_i denotes the model fit error regarding the discrepancy between the predicted value and the actual value for each observation. The data were nested within each child controlling for school.

To compare the parallel models using either fraction knowledge or grade as a predictor, we computed the Akaike information criterion (AIC) for each model (fraction knowledge AIC = 11677.69 and grade AIC = 12021.09). The resulting Δgrade = 343.53, with Akaike weight, w_{grade} = 1.61e−224, evidence ratio, ER_{grade} = 6.2e+223, and log evidence ratio, LER_{grade} = 223.79, indicates that the fraction knowledge model is far superior to the grade level model ($p < .001$; $p_i = \exp\left(\frac{11677.69-12021.09}{2}\right) < .001$; Burnham & Anderson, 2004). Furthermore, grade level was not a good proxy for fraction knowledge in our data, with wide within-grade variation (see Table 1 for distribution). Taken together with the previously discussed documented issues in using grade as a grouping variable, we do not report the grade level model.

After conducting the linear mixed model, we used STATA's *margins* command, which provides estimates of both fixed effects (the independent variables) and the effects of marginal means comparable to analysis of variance (ANOVA)-type contrasts. Significant *p* values of fixed effects and their respective interaction terms were followed up with contrasts of marginal means (Rabe-Hesketh & Skrondal, 2012; Williams, 2012). Apart from planned follow-up comparisons of format type, all other contrasts were adjusted using a Sidak correction. We report Cohen's *d* for effect sizes of pairwise comparisons by converting delta method standard errors reported by STATA's margins command into standard deviations for the denominator and using the difference between predictive marginal means as the numerator.

Results

The proportional reasoning task showed significant correlations with the fraction knowledge task across all three conditions, $r_{continuous(192)}$ = .383, $r_{discretized(182)}$ = .624, $r_{discrete(189)}$ = .479, all with values of $p < .001$. Means and standard deviation statistics for all the factors used in our crossed-effects linear mixed model are presented in Table A1 of Appendix D. The overall results of our model are summarized in Table A2 of Appendix D, which shows that gender was not a significant predictor.

Main effects of format type

All between-participant effect statistics are summarized in Table 2, and follow-up pairwise comparisons are shown in Table 3.

We expected that difficulty in solving proportional equivalence problems would be easiest for continuous formats, followed by discretized formats, with discrete formats being the hardest. There was indeed a significant format effect. Follow-up pairwise comparisons reflected differences among all three formats. As hypothesized, children in the continuous format had the highest scores, followed by the discretized format and then the discrete format. However, format differences varied for levels of fraction knowledge.

Table 1
Frequency distributions between grade and fraction knowledge.

Fraction knowledge	Grade			Total
	Third	Fourth	Fifth	
Low	95	61	43	199
Medium	58	70	57	185
High	29	63	89	181
Total	182	194	189	565

Note. $\chi^2(1, N = 565) = 53.224$, $p < .000$.

Table 2
Between-participant effects by format type.

	F	*p*	*df*
Format type	9.995	.000	2
Fraction knowledge	102.035	.000	2
Format Type * Fraction Knowledge	7.055	.007	4
Gender	1.420	.092	1

Note. .000 denotes $p < .001$.

Table 3
Overall contrasts by format type.

	F	*p*	*d*	Sidak *p*
Fraction knowledge model (all levels)				
Continuous vs. discretized	2.668	.021	.12	
Continuous vs. discrete	9.990	.000	.23	
Discretized vs. discrete	2.184	.036	.11	
Low fraction knowledge				
Continuous vs. discretized	8.487	.000	.21	.000
Continuous vs. discrete	4.930	.002	.16	.010
Discretized vs. discrete	0.442	.346	.05	1.000
Medium fraction knowledge				
Continuous vs. discretized	0.007	.904	.01	1.000
Continuous vs. discrete	5.216	.001	.17	.007
Discretized vs. discrete	4.351	.003	.15	.019
High fraction knowledge				
Continuous vs. discretized	0.014	.864	.01	1.000
Continuous vs. discrete	0.980	.161	.07	.651
Discretized vs. discrete	1.296	.108	.08	.497

Note. .000 denotes $p < .001$.

Fraction knowledge interaction with format

As expected, scores at each fraction knowledge (FK) level significantly varied by format, reflected by a significant interaction between FK and format type (see Table 2 and Fig. 2). To follow up, we examined planned contrasts of marginal means between format types within each FK level (Table 3). Follow-up contrasts show that for children with low FK, performance was significantly higher with the continuous formats than with both the discretized and discrete formats with small effects (Table 3). There were no differences between the discretized and discrete formats for low-FK children. Children with medium FK reflect a different pattern, with their scores not showing differences between the continuous and discretized formats, but both formats were significantly different from the discrete format with comparable effects (Table 3). Children in the high-FK group showed no differences among the three different format types. All other effects and interactions are reported in Appendix E.

Comparison with chance (50%) by format type at each condition

One-sample t tests were conducted to examine whether children at each format type varied from chance (50% accuracy) on each condition. We set the alpha level at .017 after a Bonferroni correction of three conditions per FK group. We found that low-FK children were no different from chance when manipulating discretized formats, $t(61) = 1.703$, $p = .094$, and discrete formats, $t(60) = 0.208$, $p = .836$, but they were significantly higher than chance with continuous formats, $t(65) = 4.422$, $p < .001$. The medium-FK group was also at chance with discrete formats, $t(56) = 2.387$, $p = .0195$. Apart from the discrete formats for the medium-FK group, the medium-FK and high-FK groups were above chance for all format types, with a range in scores for $t(56) = 5.292$ to $t(64) = 19.686$ and all $ps < .001$.

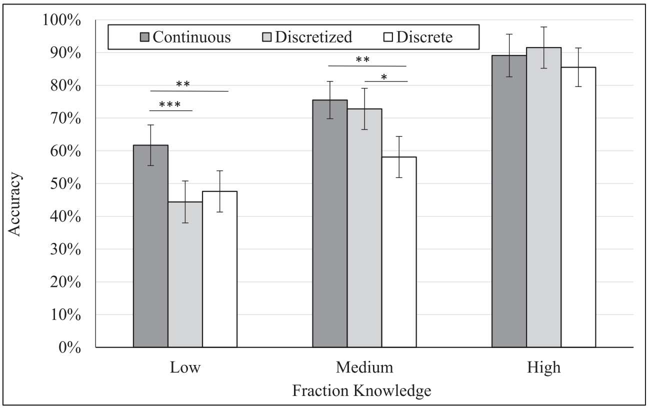


Fig. 2. Accuracy scores on the proportional equivalence task by format and fraction ability, with 50% representing chance. Error bars represent 95% confidence intervals. **p* < .05; ***p* < .01; ****p* < .005.

Discussion

The current study provides novel evidence about the effects of format, as well as scale factor and foil type, on children's equivalence judgments of visually depicted proportions. Children's performance was highest with continuous formats, followed by discretized and then discrete formats. Furthermore, children's judgments suffered as the scale factor increased as well as when trials contained a numerator-match foil (e.g., equal parts syrup between target and foil). Importantly, the effects of format, scale factor, and foil type interacted with children's level of fraction knowledge. Children with low FK performed at chance with discretized and discrete formats, whereas children with medium FK performed at chance with purely discrete formats only. These findings suggest that the perceptual nature of demarcations in discretized formats seems to disengage children's intuitive proportional reasoning skills when fraction knowledge is low, with the Gestalt of spatial separations in discrete formats exacerbating this effect. It is likely that noncontinuous formats encourage children to represent proportions using integers as they invoke counting, and it may have been the case that the Gestalt of discrete formats left children with no choice but to count, suggesting that children struggle to represent proportions as relationships between two whole numbers. It is also likely that maintaining whole numbers and thinking about their proportional relationship may have placed a greater cognitive demand than needing to reason about the proportionality of two visual analogues. Recent eye-tracking evidence with adults supports this notion; discrete displays elicited eye movements indicative of counting, whereas continuous displays seemed to encourage comparisons and magnitude estimation (Plummer, DeWolf, Bassok, Gordon, & Holyoak, 2017). This may also explain why adults were more precise and preferred discrete formats when connecting them to discrete symbolic formats of fractions (see DeWolf et al., 2015), perhaps due to formal instruction having automated counting processes and/or having improved the precision of their numerical magnitude representations. This, in turn, may have reduced the cognitive load necessary to count two sets of objects, hold in mind their numerosity, and think about the relationships between the two counted objects as a proportion.

Continuous formats are likely easier for children to process because the proportion is represented as an analogue. Support for the ease of processing analogue representations of proportions comes from data with infants when proportions are too big to count (e.g., representing a 4:1 ratio in 80:20 units) (Denison & Xu, 2010, 2014; Denison, Reed, & Xu, 2013; McCrink & Wynn, 2007; Xu & Garcia, 2008). Even at a young age, children seem to develop a generalized magnitude system that allows them to make sense of continuous intensive quantities (Hurst & Cordes, 2018; Newcombe, Levine, & Mix, 2015; Walsh, 2003). However, children often fail to use this system in situations where discrete demarcations are applied to continuous quantities (Jeong et al., 2007; Newcombe et al., 2015), and children are biased to use numerosities, which in turn leads them to erroneously represent proportions as extensive magnitudes. The discretized representations may be "in between" the continuous and discrete representations given that a child can easier disregard the demarcations and focus on the analogue representation of the proportion, which may have facilitated the encoding and maintenance of an intensive magnitude representation when compared with the discrete formats. In the current study, we found that children were less likely to be sidetracked by discretized or discrete quantities as their fraction knowledge increased. One possible explanation could be that increased fraction knowledge improves the precision of their generalized magnitude system. This idea is supported by data from the number line estimation task; estimation errors were lower for children with higher FK knowledge and vice versa.

Low- and medium-FK children's performance was negatively affected by scale factor, suggesting that mental transformations are used to enlarge/shrink proportions of varying sizes until they match—a cognitively taxing process, especially as the distance between proportions increases (Boyer & Levine, 2012; Möhring et al., 2018). This is consistent with evidence from studies on other spatial thinking tasks such as performing spatial transformations in mental rotations (Shepard & Metzler, 1971) and spatial navigation (Kosslyn, 2005), where greater angles in mental rotations and longer navigation paths increase cognitive demands and decrease performance (Kosslyn, 2005; Shepard & Metzler, 1971). A closer look shows that format further exacerbates the effect of scale factor for both low- and medium-FK children.

We also investigated whether children are systematically lured by particular foil types. As reported in the supplementary material, there was an overall effect of foil type such that children were less likely to select the correct match on trials that involved numerator foils; this effect was significant only for the low-FK children (see Table A2 and Appendix E). Importantly, and consistent with previous work (Boyer & Levine, 2012), the foils do not seem to have systematically affected most of the children in our study. Thus, the effects that we found cannot be explained by errors based on foil type exclusively.

As can be seen in Table A2 of Appendix D and Fig. A2 of Appendix E, there was an interaction of format, scale factor, foil type, and direction. As can be seen in Table A2 and Fig. A2, numerator foils seem to be harder overall, whereas for denominator foils the difficulty seems to depend on format and scaling. Our findings suggest that these differences are driven by format, with most of the differences resulting between children interacting with continuous displays versus discrete or discretized displays. The latter two formats (discretized and discrete) appear to reflect a very similar pattern. It is possible that children's performance in the continuous format on trials with denominator foils is more resistant to the effects of scaling and directionality (scaling upward vs downward) because both continuous formats and denominator foils impose fewer cognitive demands. Other unexpected results involved interactions between or among (a) fraction knowledge and scaling direction; (b) format, fraction knowledge, and magnitude; and (c) format, fraction knowledge, and foil type. The above interactions seem to be mostly driven by the differences between low-FK and med-FK children and mirror the pattern of results described in the four-way interaction above. Overall, it appears that scaling down at the higher magnitudes (.5 and .33) may be more challenging than scaling up for discretized and discrete formats. Performance detriments on discretized formats when scaling down were also reported by Boyer and Levine (2012), suggesting that the effects of scaling may be stable, but only at higher scaling magnitudes, with formats that are not intuitive and for children with frail proportional reasoning skills. Lastly, we note that our data should be interpreted with caution because our effect sizes were between small and medium, which could limit their generalization to real-world settings.

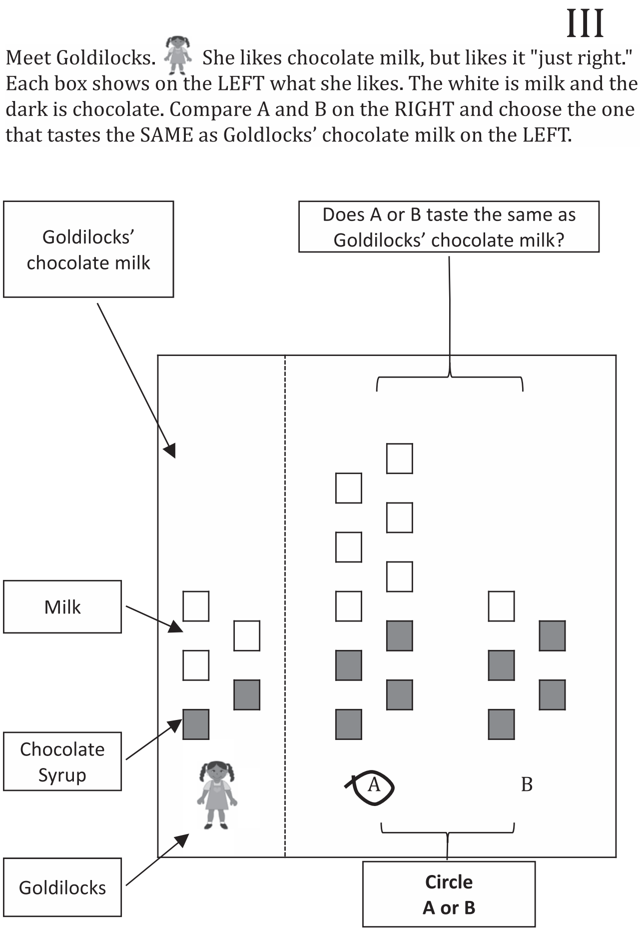
Implications

With respect to teaching practice, our findings shed light on several ways in which instruction could be sequenced. As others have suggested, proportions may be understood sooner if presented in continuous formats and lower scaling magnitudes (Boyer & Levine, 2012, 2015; Boyer et al., 2008; Goswami, 1995; Jeong et al., 2007; Singer-Freeman & Goswami, 2001; Spinillo & Bryant, 1991, 1999). Whereas recent work suggests that discrete formats are more intuitive with bipartite numerical notations (a/b) of proportions (DeWolf et al., 2015), we argue that children need significant scaffolding before they can proficiently map between symbolic and discretized/discrete nonsymbolic proportions. This may be particularly true for lower-performing children who struggle with discretized and discrete units regardless of age. The current study adds evidence providing more specificity to now familiar instructional recommendations in that the sequencing progression should move from continuous to discretized and, lastly, entirely discrete representations with gradual increases in scaling. In elucidating these distinctions, future work should experimentally investigate the optimal ways for integrating children's intuitive analogue proportional abilities with their explicit mathematical operations and concepts.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305B150014 to Temple University and by National Science Foundation (NSF) Grant SBE-1041707 to the Spatial Intelligence and Learning Center. The opinions expressed are those of the authors and do not represent the views of the Institute of Education Sciences, the Department of Education, or the NSF.

Appendix A



Appendix B

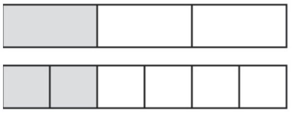
Trial	Target	Match	Foil	Foil type	Scale factor	Scaling magnitude
Scale factor > 1 (scaling up)						
1	1/4	3/12	3/4	Denominator	3:1 = 3.00	3
2	1/3	3/9	2/3	Denominator	3:1 = 3.00	3
3	2/3	6/9	2/8	Numerator	6:2 = 3.00	3
4	3/4	9/12	3/9	Numerator	9:3 = 3.00	3
5	1/4	2/8	3/4	Denominator	2:1 = 2.00	2
6	1/3	2/6	2/3	Denominator	2:1 = 2.00	2
7	2/3	4/6	2/8	Numerator	4:2 = 2.00	2
8	3/4	6/8	3/9	Numerator	6:3 = 2.00	2
9	2/8	3/12	2/3	Numerator	3:2 = 1.50	1
10	2/6	3/9	2/3	Numerator	3:2 = 1.50	1
11	4/6	6/9	2/6	Denominator	6:4 = 1.50	1
12	6/8	9/12	2/8	Denominator	9:6 = 1.50	1
Scale factor < 1 (scaling down)						
13	3/12	2/8	3/4	Numerator	2:3 = 0.67	1
14	3/9	2/6	3/4	Numerator	2:3 = 0.67	1
15	6/9	4/6	3/9	Denominator	4:6 = 0.67	1
16	9/12	6/8	4/12	Denominator	6:9 = 0.67	1
17	2/8	1/4	2/3	Numerator	1:2 = 0.50	2
18	2/6	1/3	4/6	Denominator	1:2 = 0.50	2
19	4/6	2/3	4/12	Numerator	2:4 = 0.50	2
20	6/8	3/4	2/8	Denominator	3:6 = 0.50	2
21	3/12	1/4	3/4	Numerator	1:3 = 0.33	3
22	3/9	1/3	3/4	Numerator	1:3 = 0.33	3
23	6/9	2/3	3/9	Denominator	2:6 = 0.33	3
24	9/12	3/4	4/12	Denominator	3:9 = 0.33	3

Note. We used the same items as Boyer and Levine (2012). However, the order was randomized such that items were presented to every child in every format type in the following trial order: 1, 13, 6, 22, 10, 8, 24, 15, 19, 20, 3, 12, 16, 17, 9, 5, 21, 4, 2, 11, 7, 18, 23, 14.
Source. Table taken from Boyer and Levine (2012).

Appendix C

The figures below represent two fractions. Which number sentences represent the addition and sum of these fractions? Circle yes or no.

1. The figures below represent two fractions. Which number sentences represent the addition and sum of these fractions? Circle yes or no.



a. $\frac{1}{2} + \frac{2}{4} = \frac{4}{4}$ yes no

b. $\frac{1}{3} + \frac{2}{6} = \frac{4}{6}$ yes no

c. $\frac{1}{3} + \frac{2}{6} = \frac{3}{18}$ yes no

d. $\frac{1}{3} + \frac{2}{6} = \frac{3}{9}$ yes no

e. $\frac{2}{6} + \frac{2}{6} = \frac{4}{6}$ yes no

2. Compare each fraction to $\frac{4}{12}$ without performing any calculations. Circle larger or smaller.

a. $\frac{4}{10}$ larger smaller

b. $\frac{4}{16}$ larger smaller

c. $\frac{2}{12}$ larger smaller

d. $\frac{8}{12}$ larger smaller

3. Celia filled 1 of her buckets with candy. Her friends ate half of it. Which fractions represent the amount of candy that Celia had left?
Circle yes or no.

a. $\frac{3}{6}$ yes no

b. $\frac{1}{2}$ yes no

c. $\frac{1}{4}$ yes no

d. $\frac{1}{1}$ yes no

e. $\frac{2}{2}$ yes no

4. Victoria has $\frac{3}{4}$ of a liter of milk which she wishes to split into cups, each $\frac{1}{4}$ of a liter. How many cups will she have?

a. 3

b. 4

c. 1

d. 8

Appendix D

See Tables A1–A3.

Table A1
Summary of proportional equivalence task of overall accuracy, scaling magnitude, foil type, and scaling direction by condition, grade level, and fraction knowledge.

	Condition			Fraction knowledge			Grade level		
	Continuous	Discretized	Discrete	Low	Medium	High	Third	Fourth	Fifth
Overall accuracy	74	69	65	51	69	87	63	71	74
	(.27)	(.31)	(.29)	(.26)	(.30)	(.21)	(.28)	(.30)	(.28)
Continuous	–	–	–	63	72	90	68	76	78
	–	–	–	(.25)	(.28)	(.21)	(.28)	(.28)	(.25)
Discretized	–	–	–	44	71	91	64	70	73
	–	–	–	(.25)	(.29)	(.17)	(.30)	(.31)	(.32)
Discrete	–	–	–	50	59	84	56	65	72
	–	–	–	(.24)	(.31)	(.21)	(.26)	(.31)	(.29)
Magnitude 1 (.33/1.5)	78	73	71	62	73	86	71	74	77
	(.26)	(.29)	(.28)	(.26)	(.29)	(.23)	(.26)	(.29)	(.28)
Magnitude 2 (.50/2)	76	69	64	51	70	88	62	71	76
	(.27)	(.32)	(.32)	(.27)	(.31)	(.21)	(.30)	(.32)	(.28)
Magnitude 3 (.67/3)	70	64	58	41	64	86	54	67	71
	(.32)	(.37)	(.37)	(.32)	(.36)	(.24)	(.36)	(.36)	(.34)
Denominator foil	82	70	64	56	72	87	64	73	78
	(.25)	(.32)	(.31)	(.29)	(.30)	(.23)	(.31)	(.31)	(.28)
Numerator foil	68	68	65	47	66	87	61	68	72
	(.35)	(.33)	(.33)	(.31)	(.34)	(.22)	(.33)	(.34)	(.32)
Scaling up	75	70	66	53	70	87	64	71	76
	(.28)	(.32)	(.32)	(.28)	(.31)	(.23)	(.31)	(.31)	(.29)
Scaling down	75	68	64	50	68	87	62	70	74
	(.28)	(.31)	(.30)	(.26)	(.30)	(.20)	(.29)	(.30)	(.29)

Note. Values are percentages (%). Standard deviations are reported in parentheses.

Table A2
Summary results of crossed-random-effects linear mixed model.

	F	*p*	*df*
Main effects			
Within participants			
Scale factor	90.340	.000	2
Foil type	31.330	.000	1
Scale direction	2.860	.017	1
Two-way interactions			
Scale Factor * Fraction Knowledge	79.025	.000	4
Foil Type * Format Type	52.160	.000	2
Foil Type * Fraction Knowledge	14.575	.000	2
Scale Direction * Fraction Knowledge	5.340	.005	2
Scale Factor * Scale Direction	9.475	.000	2
Scale Factor * Foil Type	6.020	.002	2
Three- and four-way interactions			
Scale Factor * Fraction Knowledge * Format Type	11.360	.004	8
Scale Factor * Foil Type * Format Type	6.855	.008	4
Foil Type * Fraction Knowledge * Format Type	10.065	.001	4
Foil Type * Scale Direction * Format Type	4.285	.014	2
Scale Factor * Foil Type * Scale Direction * Format Type	7.720	.004	4

Note. .000 denotes *ps* < .001. All other interactions were not significant (2.842 < *F*s < .142, .058 < *ps* < .961). The *df*s for error terms ranged from 553.17 to 554.08.

Table A3
Summary of contrasts of marginal means between all format types at each scaling factor and scaling direction.

	F	Sidak *p*	*d*
Scaling Factor 1 × Denominator Foil × Up			
Continuous vs Discrete	15.961	.000	.29
Scaling Factor 1 × Denominator Foil × Down			
Continuous vs. discretized	5.917	.014	.18
Continuous vs. discrete	9.202	.000	.22
Scaling Factor 2 × Denominator Foil × Up			
Continuous vs. discretized	9.159	.000	.22
Continuous vs. discrete	7.144	.004	.19
Scaling Factor 2 × Denominator Foil × Down			
Continuous vs. discretized	6.771	.006	.22
Continuous vs. discrete	18.973	.000	.19
Scaling Factor 2 × Numerator Foil × Up			
Discrete vs. discretized	6.552	.007	.19
Scaling Factor 3 × Denominator Foil × Down			
Continuous vs. discretized	7.960	.002	.21
Continuous vs. discrete	29.338	.000	.40
Discretized vs. discrete	6.372	.008	.19
Scaling Factor 3 × Denominator Foil × Up			
Continuous vs. discrete	6.372	.008	.18

Note. Only contrasts with alpha below .05 after the Sidak adjustment are shown.
.000 denotes *ps* < .001. The *df*s for error terms are as follows: continuous vs. discretized = 366, continuous vs. discrete = 372, and discretized vs. discrete = 368.

Appendix E

Results of interaction terms

Fraction knowledge interaction with scaling direction

Fraction knowledge also significantly interacted with (a) scaling direction, (b) scaling magnitude, and (c) foil type (see Table A2 of Appendix D). Follow-up contrasts at each fraction knowledge with adjusted p values using the Sidak method suggest that children with low fraction knowledge (FK) perform significantly worse when scaling down than when scaling up, $F(1, 196) = 16.565$, Sidak $p < .001$, $d = .29$, but there were no significant differences between scaling direction within the medium-FK children, $F(1, 182) = 0.022$, Sidak $p = .995$, or high-FK children, $F(1, 178) = 0.001$, Sidak $p = 1.000$ (see Table A1 of Appendix D for means and standard deviations). The two-way interactions, Fraction Knowledge × Scaling Factor and Fraction Knowledge × Foil Type, are subsumed within the three-way interactions of Scaling Factor × Format × Fraction Knowledge and Foil Type × Fraction Knowledge × Format (Fig. A1) (discussed in the next section).

Fraction knowledge interaction with scaling factor and format

Fraction knowledge interacted with scaling magnitude and format (see Table A2 of Appendix D and Fig. A1). Contrasts of marginal effects between format types at each FK level with Sidak adjustments suggest that low-FK children performed significantly better with continuous formats than with discretized formats, $F_{mag2}(1, 128) = 11.713$, Sidak $p < .001$, $d = .418$; $F_{mag3}(1, 128) = 7.801$, $p = .001$, $d = .341$, or with discrete formats, $F_{mag2}(1, 128) = 8.736$, Sidak $p = .001$, $d = .361$, $F_{mag3}(1, 128) = 6.160$, Sidak $p = .008$, $d = .303$, at Magnitudes 2 and 3, respectively. Medium-FK children performed better with continuous formats than with discrete formats at Magnitude 2, $F_{mag2}(1, 120) = 4.621$, Sidak $p = .041$, $d = .271$, and Magnitude 3, $F_{mag3}(1, 125) = 5.216$, Sidak $p = .022$, $d = .282$. The remaining contrasts were not significantly different.

Fraction knowledge interaction with foil type and format type

Fraction knowledge also interacted with foil type and format type (see Table A2 of Appendix D and Fig. A2). Follow-up contrasts of marginal means with Sidak adjustments suggest that low-FK children had better accuracy with continuous formats than with discretized formats, $F_{Den}(1, 129) = 18.000$, $p < .001$, $d = .518$, and discrete formats, $F_{Den}(1, 129) = 12.301$, $p < .001$, $d = .429$, for denominator foils.

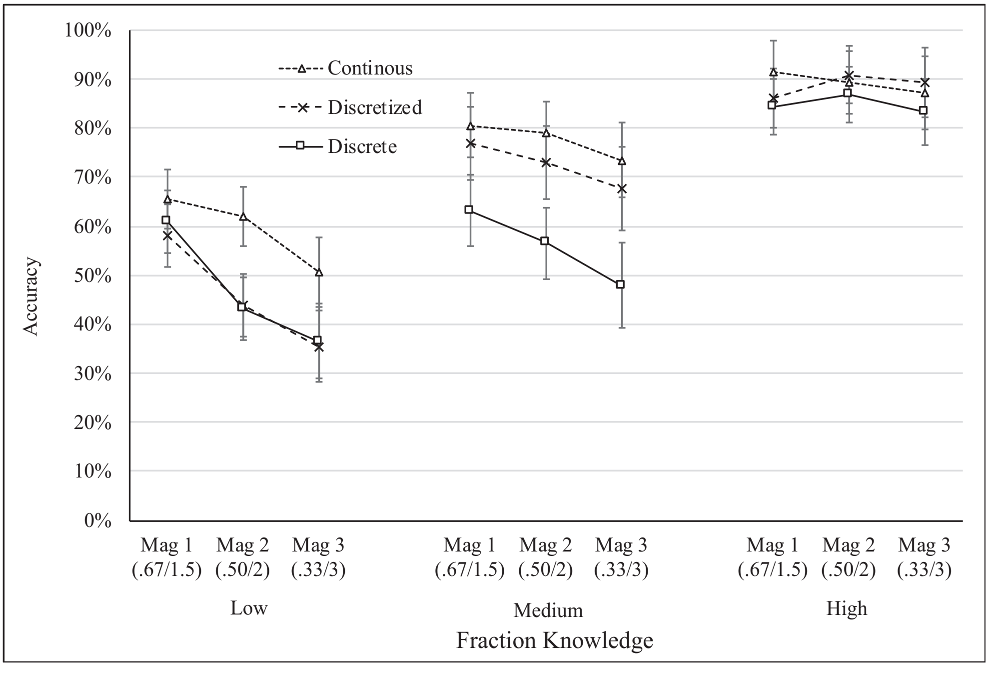


Fig. A1. Proportional equivalence scores at each scaling magnitude (Mag) by fraction knowledge and format type. Error bars represent 95% confidence intervals.

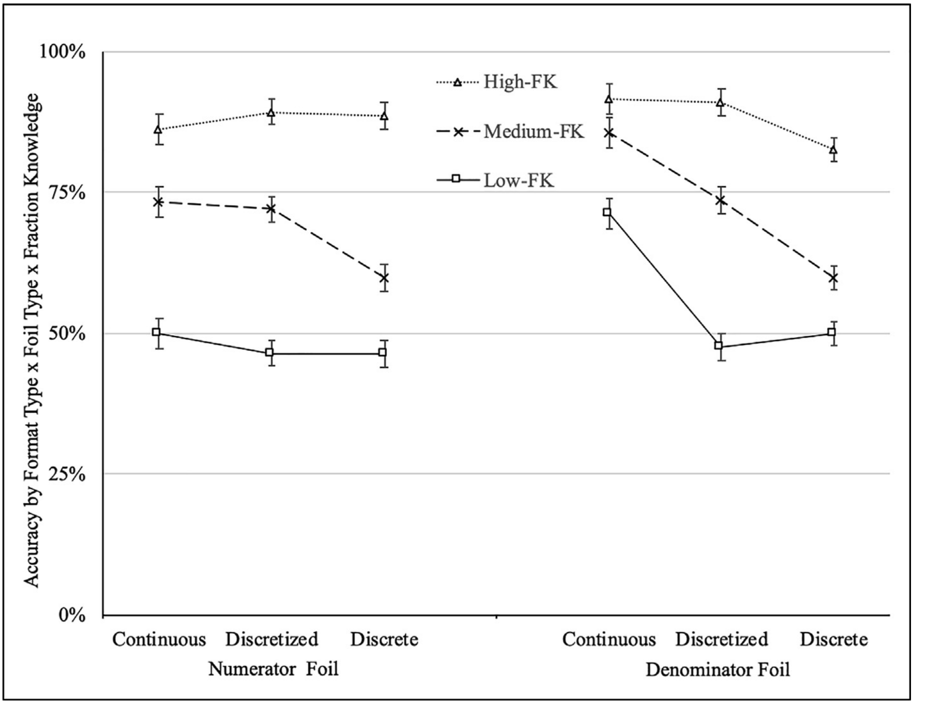


Fig. A2. Proportional equivalence scores at each format type by fraction knowledge (FK) and foil type. Error bars represent 95% confidence intervals.

Medium-FK children performed better with continuous formats than with discrete formats for denominator foils, $F_{Den}(1, 126) = 11.281$, $p < .001$, $d = .415$. All other contrasts were not significantly different.

Scale factor, scaling direction, foil type, and format interaction

There was a significant interaction among scale factor, scale direction, foil type, and format. As shown in Fig. A3, the difficulty of denominator foils increases with higher scaling magnitude, especially when scaling down, with the greatest differences being between the continuous formats and discrete formats at the highest magnitude scaled down. To unpack the contribution of each factor, we contrast marginal means with format as a between-participant factor at each of the six scaling magnitudes (three magnitudes up and three magnitudes down). We report pairwise comparisons when the contrasts revealed a significant effect of format. (A summary of the contrasts with significant effects of format and follow-up pairwise comparisons with Sidak adjustments is included in Table A3 of Appendix D.)

Children performed better with continuous formats than with discrete formats at all scaling magnitudes on both scaling directions for denominator foils, with the exception of scaling up at Magnitudes 1 and 3. Children performed better with continuous formats than with discretized formats at Magnitude 1 when scaling down, Magnitude 2 on both scaling directions, and Magnitude 3 when scaling down. Children were also more accurate with discretized formats than with discrete formats at Magnitude 3 for denominator foils scaled down. For numerator foils, children performed worse with discrete formats only when scaling down at Magnitude 2. All other comparisons were not significant after the Sidak adjustment or were above .05. Lastly, all other effects and interactions were not significant.

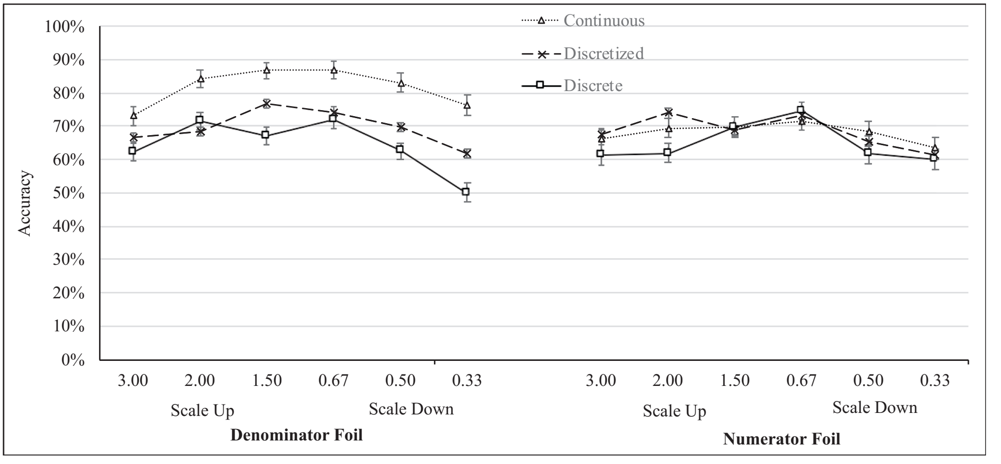


Fig. A3. Proportional equivalence scores at each scaling magnitude and direction. Error bars represent 95% confidence intervals.

References

Barth, H., Baron, A., Spelke, E., & Carey, S. (2009). Children's multiplicative transformations of discrete and continuous quantities. *Journal of Experimental Child Psychology, 103*, 441–454.

Begolli, K. N., & Richland, L. E. (2016). Teaching mathematics by comparison: Analog visibility as a double-edged sword. *Journal of Educational Psychology, 108*, 194–213.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*, 189–201.

Boyer, T. W., & Levine, S. C. (2012). Child proportional scaling: Is 1/3 = 2/6 = 3/9 = 4/12? *Journal of Experimental Child Psychology, 111*, 516–533.

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

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

Brown, G., & Quinn, R. J. (2006). Algebra students' difficulty with fractions: An error analysis. *The Australian Mathematics Teacher, 62*(4), 28–40.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*, 261–304.

Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science, 10*, 408–411.

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*, 243–249.

Denison, S., & Xu, F. (2010). Integrating physical constraints in statistical inference by 11-month-old infants. *Cognitive Science, 34*, 885–908.

Denison, S., & Xu, F. (2014). The origins of probabilistic inference in human infants. *Cognition, 130*, 335–347.

DeWolf, M., Bassok, M., & Holyoak, K. J. (2015). Conceptual structure and the procedural affordances of rational numbers: Relational reasoning with fractions and decimals. *Journal of Experimental Psychology: General, 144*, 127–150.

Fischbein, E., & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? *Educational Studies in Mathematics, 15*(1), 1–24.

Fryer, R. G., Jr., & Levitt, S. D. (2004). Understanding the black–white test score gap in the first two years of school. *Review of Economics and Statistics, 86*, 447–464.

Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., ... Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology, 105*, 683–700.

Goswami, U. (1995). Transitive relational mappings in three- and four-year-olds: The analogy of Goldilocks and the Three Bears. *Child Development, 66*, 877–892.

Hansen, N., Jordan, N. C., Fernandez, E., Siegler, R. S., Fuchs, L., Gersten, R., & Micklos, D. (2015). General and math-specific predictors of sixth-graders' knowledge of fractions. *Cognitive Development, 35*, 34–49. https://doi.org/10.1016/j.cogdev.2015.02.001.

Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth, 17*, 267–321.

Hecht, S. A., & Vagi, K. J. (2010). Sources of group and individual differences in emerging fraction skills. *Journal of Educational Psychology, 102*, 843–859.

Hecht, S. A., & Vagi, K. J. (2012). Patterns of strengths and weaknesses in children's knowledge about fractions. *Journal of Experimental Child Psychology, 111*, 212–229.

Hoffrage, U., Gigerenzer, G., Krauss, S., & Martignon, L. (2002). Representation facilitates reasoning: What natural frequencies are and what they are not. *Cognition, 84*, 343–352.
Howe, C., Nunes, T., & Bryant, P. E. (2011). Rational number and proportional reasoning: Using intensive quantities to promote achievement in mathematics and science. *International Journal of Science and Mathematics Education, 9*, 391–417.
Hurst, M. A., & Cordes, S. (2018). Attending to relations: Proportional reasoning in 3- to 6-year-old children. *Developmental Psychology, 54*, 428–439.
Jäger, S., & Wilkening, F. (2001). Development of cognitive averaging: When light and light make dark. *Journal of Experimental Child Psychology, 79*, 323–345.
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.
Jirout, J., & Newcombe, N. S., (2019). Relative magnitude as a key idea in mathematics cognition. In K. S. Mix & M. Battista (Eds.), *Visualizing mathematics: The role of spatial reasoning in mathematical thought* (pp. 3–24). New York: Springer.
Kosslyn, S. M. (2005). Mental images and the brain. *Cognitive Neuropsychology, 22*, 333–347.
Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017). From "sense of number" to "sense of magnitude": The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences, 40*, e164.
Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychological Science, 16*, 841–845.
Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General, 111*, 1–22.
Mazzocco, M. M. M., & Devlin, K. T. (2008). Parts and "holes": Gaps in rational number sense among children with vs. without mathematical learning disabilities. *Developmental Science, 11*, 681–691.
McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science, 18*, 740–745.
Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: Is number one of them? *Psychological Bulletin, 128*, 278–294.
Möhring, W., Frick, A., & Newcombe, N. S. (2018). Spatial scaling, proportional thinking, and numerical understanding in 5- to 7-year-old children. *Cognitive Development, 45*, 57–67.
Möhring, W., Newcombe, N. S., & Frick, A. (2014). Zooming in on spatial scaling: Preschool children and adults use mental transformations to scale spaces. *Developmental Psychology, 50*, 1614–1619.
Möhring, W., Newcombe, N. S., & Frick, A. (2015). The relation between spatial thinking and proportional reasoning in preschoolers. *Journal of Experimental Child Psychology, 132*, 213–220.
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.
National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics* (Vol. 1) Reston, VA: Author.
National Mathematics Advisory Panel (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel* (Vol. 37) Washington, DC: Author http://www2.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf.
Newcombe, N. S., Levine, S. C., & Mix, K. S. (2015). Thinking about quantity: The intertwined development of spatial and numerical cognition. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*, 491–505.
Newcombe, N. S., Möhring, W., & Frick, A. (2018). How big is many? Development of spatial and numerical magnitude understanding. In A. Henik & W. Fias (Eds.), *Heterogeneity of function in numerical cognition* (pp. 157–176). New York: Elsevier.
Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology, 55*, 169–195.
Paik, J. H., & Mix, K. S. (2003). US and Korean children's comprehension of fraction names: A reexamination of cross-national differences. *Child Development, 74*, 144–154.
Piaget, J., & Inhelder, B. (1974). *The child's conception of quantities*. London: Routledge & Kegan Paul.
Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children*. New York: Norton.
Plummer, P., DeWolf, M., Bassok, M., Gordon, P. C., & Holyoak, K. J. (2017). Reasoning strategies with rational numbers revealed by eye tracking. *Attention, Perception, & Psychophysics, 79*, 1426–1437.
Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata* (3rd ed.). College Station, TX: Stata Press.
Schumacher, R. F., Namkung, J. M., Malone, A., & Fuchs, L. S. (2012). Fraction battery. (Available from L. S. Fuchs, 228 Peabody, Vanderbilt University, Nashville, TN 37203, USA).
Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701–703.
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*, 13–19.
Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*, 273–296.
Singer-Freeman, K. E., & Goswami, U. (2001). Does half a pizza equal half a box of chocolates? Proportional matching in an analogy task. *Cognitive Development, 16*, 811–829.
Spinillo, A. G., & Bryant, P. E. (1991). Children's proportional judgments: The importance of "half". *Child Development, 62*, 427–440.
Spinillo, A. G., & Bryant, P. E. (1999). Proportional reasoning in young children: Part–part comparisons about continuous and discontinuous quantity. *Mathematical Cognition, 5*, 181–197.
Tian, J., & Siegler, R. S. (2018). Which type of rational numbers should students learn first? *Educational Psychology Review, 30*, 351–372.
Van Dooren, W., Lehtinen, E., & Verschaffel, L. (2015). Unraveling the gap between natural and rational numbers. *Learning and Instruction, 37*, 1–4.
Vasilyeva, M., Duffy, S., & Huttenlocher, J. (2007). Developmental changes in the use of absolute and relative information: The case of spatial extent. *Journal of Cognition and Development, 8*, 455–471.
Vasilyeva, M., & Huttenlocher, J. (2004). Early development of scaling ability. *Developmental Psychology, 40*, 682–690.

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Science, 7*, 483–488.
Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20*, 273–281.
Williams, R. (2012). Using the margins command to estimate and interpret adjusted predictions and marginal effects. *The Stata Journal, 12*(2), 308–331.
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.
Zhu, L., & Gigerenzer, G. (2006). Children can solve Bayesian problems: The role of representation in mental computation. *Cognition, 98*, 287–308.