Developmental Psychology
2017, Vol. 53, No. 3, 587–596

0012-1649/17/$12.00 http://dx.doi.org/10.1037/dev0000252

BRIEF REPORT

The Number Line Is a Critical Spatial-Numerical Representation: Evidence From a Fraction Intervention

Noora Hamdan and Elizabeth A. Gunderson
Temple University

Children's ability to place fractions on a number line strongly correlates with math achievement. But does the number line play a causal role in fraction learning or does it simply index more advanced fraction knowledge? The number line may be a particularly effective representation for fraction learning because its properties align with the desired mental representation and take advantage of preexisting spatial-numeric biases. Using a pretest-training-posttest design, we examined second and third graders' fraction learning in 3 conditions: number line training, area model training, and a non-numerical control. Children who received number line training improved at representing fractions with number lines, and children who received area model training improved at representing fractions with area models. Critically, only the number line training led to transfer to an untrained fraction magnitude comparison task. We conclude that the number line plays a causal role in children's fraction magnitude understanding, and is more beneficial than the widely used area model.

Keywords: fractions, area model, number line, math cognition

Supplemental materials: http://dx.doi.org/10.1037/dev0000252.supp

Fraction knowledge is crucial for math achievement (e.g., Booth & Newton, 2012; Siegler et al., 2012). Fractions are U.S. children's first introduction to a number system that extends whole numbers to include rational and real numbers (Institute of Education Sciences, 2010), and require reorganization of number understanding to accommodate the infinite set of numbers that exist between whole numbers (Siegler & Pyke, 2013). With fraction arithmetic, children learn that properties of whole numbers are not properties of all numbers (Ni & Zhou, 2005; Vamvakoussi & Vosniadou, 2010). For example, multiplying two whole numbers always leads to a larger number than either operand, but this is not true for fractions. Thus, fractions pose distinct challenges for students, while being uniquely important for math achievement (Siegler, Thompson, & Schneider, 2011).

The critical role of fraction magnitude knowledge in mathematics is borne out empirically. Middle school fraction magnitude estimation, measured by number line estimation of unit fractions (fractions with a numerator of 1), predicts algebra readiness, learning, and mastery (Booth & Newton, 2012). Eighth-graders' fraction magnitude comparison and estimation accounts for a larger portion of variance in mathematics achievement tests than fraction arithmetic (Siegler et al., 2011). Thus, fraction magnitude understanding is a crucial milestone in mathematical development and a gatekeeper for subsequent math achievement (Booth & Newton, 2012).

Unfortunately, many students do not pass that gate. Despite the predictive power of fraction magnitude knowledge for math achievement, students experience challenges with basic fraction concepts. For example, only 50% of U.S. eighth graders can correctly order the magnitudes of the fractions [2/7], [1/12], and [5/9] from least to greatest (National Council of Teachers of Mathematics, 2006). We ask whether typical early fraction instruction exacerbates this problem.

In U.S. schools, fractions are introduced in the early elementary years predominately using area models—two-dimensional canonical shapes that children divide into parts (Common Core State Standards Initiative, 2015). In third grade, Common Core goals for fraction learning include comparing two fractions, ordering fractions, reasoning about their amounts, and demonstrating understanding of equivalence (Common Core State Standards Initiative, 2015). Although the Common Core encourages use of fraction number lines starting in third grade, the focus on area models prior to this raises concern that fractions will be viewed primarily as part–whole relations (Opfer & Siegler, 2012). In working with a circular pizza, for example, a child might fixate on the number of pieces of pizza to be eaten (the numerator) and disregard the total

This article was published Online First November 17, 2016.

Noora Hamdan and Elizabeth A. Gunderson, Department of Psychology, Temple University.

This study was supported by National Science Foundation (NSF) CAREER DRL-1452000 and NSF SBE-1041707 (Spatial Intelligence and Learning Center). We thank all of the principals, teachers, and students who gave their time to this research, and the research assistants and interns who helped to conduct it: Alaina Chlebek, Alexander D'Esterre, Mahala Femovich. Courtney Gray, Marisol Savage, and Brooke Singer. We also thank Nora Newcombe for helpful feedback on a previous draft.

Correspondence concerning this article should be addressed to Elizabeth A. Gunderson, Department of Psychology, Temple University, 1701 North 13th Street, Philadelphia, PA 19122. E-mail: liz.gunderson@temple.edu

number of pieces (the denominator). Further, two-dimensional area models do not align with the unidimensionality of the real number system (Institute of Education Studies, 2010). This may contribute to a lack of understanding that fractions are fundamentally magnitudes (Newcombe, Levine, & Mix, 2015).

The Number Line Representation

Given these problems with the area model, we examine the benefits of a different representation, the number line, which has several properties that align with fraction magnitude concepts. The unidimensionality of the number line captures the fact that fractions are continuous number magnitudes that can be ordered with whole numbers and other rational and real numbers on one dimension (Siegler & Lortie-Forgues, 2014). The number line also takes advantage of preexisting spatial–numeric associations: the association between numerical magnitude and line length (de Hevia & Spelke, 2010; Lourenco & Longo, 2010; Newcombe et al., 2015), the association of small numbers with the left side of space and large numbers with the right side of space (Dehaene, Bossini, & Giraux, 1993; Patro & Haman, 2012), and the ability to match nonsymbolic fractional quantities based on spatial extent (Boyer & Levine, 2015). Thus, the number line may be particularly beneficial for teaching fraction magnitudes, providing a powerful link between spatial extent and fraction magnitude (Gunderson, Ramirez, Beilock, & Levine, 2012; LeFevre et al., 2013; Siegler & Lortie-Forgues, 2014).

The benefits of the number line have been well studied for whole numbers. High levels of linearity and low estimation error when placing whole numbers on a number line are related to math skills in both correlational (e.g., Geary, Hoard, Nugent, & Byrd-Craven, 2008; Holloway & Ansari, 2009; Schneider, Grabner, & Paetsch, 2009) and experimental (Booth & Siegler, 2008) studies. These benefits have been shown in children as early as preschool (e.g., Ramani & Siegler, 2008), and on skills including approximate calculation (Booth & Siegler, 2008; Gunderson et al., 2012), counting ability (Hoffmann, Hornung, Martin, & Schiltz, 2013), and number comparison (Fazio, Bailey, Thompson, & Siegler, 2014).

Although the benefit of number line knowledge for whole numbers is clear, research concerning its role in fraction magnitude understanding is mostly correlational. The evidence that does exist indicates that magnitude knowledge is crucial for a deeper understanding of fractions (Siegler et al., 2011). For example, evidence from U.S. and international samples shows that fraction number line estimation predicts later fraction arithmetic, algebra skill, and math achievement (Bailey, Hoard, Nugent, & Geary, 2012; Bailey et al., 2015; Booth & Newton, 2012; Siegler et al., 2012; Siegler & Pyke, 2013; Torbeyns, Schneider, Xin, & Siegler, 2015).

Despite these strong correlations, no study has established direct causal evidence of the number line's importance for fraction learning. Although one experimental study found that emphasizing fractions on a number line as part of a larger curriculum benefited at-risk fourth graders more than the standard curriculum (Fuchs et al., 2013), the design of this prior study (in which the experimental and standard curricula differed on many dimensions other than the use of the number line) did not attempt to *isolate* the number line's role in improving children's fraction magnitude concepts. We aim to fill this gap.

The Present Study

We used a pretest-training-posttest design to rigorously test the causal effect of the number line on fraction magnitude knowledge, comparing a number line training to a well-matched area model training and a non-numerical control. Although previous work focused on older children with some fraction knowledge (fourth graders and above), we asked whether the number line would be more beneficial than the area model at the start of fraction learning, in second and third grades. We predicted that both training groups would improve on the representation taught. Importantly, we predicted that only children who received the number line training would transfer their knowledge to an untrained fraction magnitude comparison task. We defined transfer as improvement in performance on a task that was not trained (e.g., Braithwaite & Goldstone, 2013). Comparing performance across conditions on this transfer task is critical to determine whether children merely learned a procedural rule about how to segment and label the number line or area model, or whether they gained a more conceptual understanding of fraction magnitudes.

Our primary outcomes of interest are number line estimation and magnitude comparison, as these measures predict later fraction arithmetic and math achievement from fifth to eighth grades (e.g., Siegler et al., 2011). Because children in this study have minimal fraction experience, we focus on these foundational skills.

Method

Participants

A total of 114 children (69 females) in second and third grades participated, with an average age of 8.55 years (SD = .58). Children were recruited from five Catholic schools in a large U.S. city. Schools followed the Common Core State Standards (2015). Demographic information was available at the school level. Between 37% and 56% of students in each school were eligible for free or reduced lunch. On average, 67.0% of students were African American (range across schools = 3% to 99.5%), 23.9% were Caucasian (range = 0.5% to 97.0%), 4.7% were Asian (range = 0.0% to 10.0%), 1.9% were Hispanic (range = 0.0% to 4.3%), and 2.8% were other or multiple races (range = 0.0% to 7.5%).

Procedure

We used a pretest-training-posttest design. Two 30-min sessions were conducted on separate days, an average of 3.64 days apart (SD = 2.77, range = 1 to 14 days).[1] In Session 1, children completed pretest measures in one of two orders (number line estimation, magnitude comparison, area model estimation, or the

[1] For the majority of children (n = 103), the range of days between sessions was 7 days or fewer. Although we attempted to keep the range of days for all students similar, and less than a week, for 11 children, the range of days between sessions was 8 to 14 days. The reason for this greater range of days between sessions for some children included children being repeatedly absent from school or children being involved in other school activities. However, the range of days between sessions did not moderate the impact of condition on our three main outcome measures (number line PAE, area model PAE, and magnitude comparison accuracy).

reverse).[2] In Session 2, children completed the reading achievement measure, and then one of three conditions, randomly assigned within classroom: number line training (n = 39), area model training (n = 37), or crossword puzzle control (n = 38). Children then completed posttests (area model estimation, number line estimation, and magnitude comparison) in the opposite order as in Session 1.

Training Conditions

Number line training. The number line used in training was a thin horizontal rectangle with endpoints from 0 to 1 (see Figure 1 for training overview; see the online supplemental materials for the training script). We used a slightly two-dimensional number line during training to avoid a common error children make with ruler measurement: counting hatch marks as objects instead of counting the spaces between hatch marks (e.g., Solomon, Vasilyeva, Huttenlocher, & Levine, 2015). We expected that using a thin rectangle would guide children's attention to the spaces between hatch marks as the meaningful units.

First, the experimenter introduced the number line and explained how it can represent fractions. Next, the experimenter explained how to interpret the numerator and denominator to segment and shade the number line. Finally, the experimenter placed the fraction on the number line with a hatch-mark, writing the fraction above the hatch mark. Experimenters demonstrated this procedure for the unit fractions [1/2], [1/4], and [1/5].

After each experimenter demonstration, children were given a presegmented number line, and were asked to shade the correct number of segments, place the fraction by drawing a hatch mark, and write the fraction above the hatch mark. Given their young age and lack of fraction experience, children began with this segmented number line for scaffolding purposes. Afterward, children were presented with an unsegmented number line, and asked to segment, shade, and place the fraction at the correct location. After each step, children were shown the correct answer for that step of the procedure (segment, shade, and place). If successful, they moved to the next step. If unsuccessful, they were given one opportunity to redo that step. After children completed these steps for all fractions with a given denominator, the experimenter demonstrated the procedure for the unit fraction of the next denominator, and the procedure started again. Trained fractions were [1/2], [1/4], [2/4], [3/4], [1/5], [2/5], [3/5], and [4/5]. Training lasted about 15 min.

Area model training. Area model training paralleled the number line training (see Figure 2). Two exceptions reflect inherent differences between these representations: There was no mention of endpoints, and children were not required to begin shading from a certain location, whereas starting from the left endpoint (zero) was crucial to the number line training.

Crossword condition. The crossword condition included a set of three age-appropriate crossword puzzles, timed for 15 min to match the time spent on fraction training. The child worked one-on-one with the experimenter. The experimenter demonstrated solving the first word, and children were instructed to read subsequent clues aloud and match them to the words provided.

Measures

Fraction pretest and posttest measures. Figure 3 provides examples of each measure.

Number line estimation. The number line estimation task was similar to that used by others (e.g., Siegler et al., 2011) so that our findings could be compared with those studies. Its validity is observed in correlations between performance on this task and fraction arithmetic, algebra skill, and math achievement, at least through middle school (e.g., Bailey et al., 2012, 2015; Siegler & Pyke, 2013; Torbeyns et al., 2015). Further, pre- and posttest number line performance were significantly correlated, $r(110) = 0.33$, $p < .001$.

Fourteen number lines were individually presented on sheets of paper. Each number line was 22 cm long and labeled "0" at the left and "1" at the right. Above each line was the fraction to be estimated. The experimenter explained, "This is a number line. How can you show [1/2]? To show this, you need to draw a line like this." The experimenter placed a hatch mark at [1/2]. To encourage children to utilize the entire line, the experimenter stated, "When you are answering these questions, you can make a mark anywhere on the line that you think is right." Children responded by placing a hatch mark on the number line. The numbers estimated included trained (halves, fourths, and fifths) and untrained (thirds, sixths, and sevenths) fractions, presented in one of two pseudorandom orders (see the online supplemental materials). Half of the items were below [1/2] and half were at or above [1/2] to avoid bias toward either endpoint. Percent absolute error (PAE) scores were calculated for each item (PAE = |estimated line length percentage − correct percentage|), and then averaged. Lower PAE scores indicate higher accuracy. The number line estimation task showed adequate reliability (Cronbach's $\alpha_{pretest} = .49$; $\alpha_{posttest} = .78$).

Area model estimation. The area model estimation task paralleled the number line estimation task (see the online supplemental materials). This is a valid measure for assessing children's fraction estimation, as it is similar to the extensive use of shapes in school math curricula for introducing fractions (Common Core State Standards Initiative, 2015). Further, pre- and posttest area model estimation were significantly correlated, $r(111) = 0.36$, $p < .001$.

Fourteen circles were individually presented on sheets of paper; each circle was 7 cm in diameter. Above each circle was the fraction to be estimated. The experimenter explained the goal of this task: "This is a whole circle. How can you show [1/2]? To show this, you need to color part of the circle like this." The experimenter then shaded half of the circle. To encourage children to utilize the entire circle, the experimenter stated, "When you are answering these questions, you can color as much of the circle that you think is right." Children responded by shading the estimated

[2] At pretest, we also administered measures assessing children's spatial skills (proportional reasoning, spatial working memory, and mental rotation), in order to examine whether spatial skills moderated the impact of condition on fraction knowledge. For half of the participants, these measures were completed before the fraction pretests, and for the other half, they were completed after the fraction pretests. These spatial skill measures did not separately nor compositely moderate the impact of condition on our key outcomes (number line PAE, area model PAE, and magnitude comparison), and are therefore not included in our analyses.

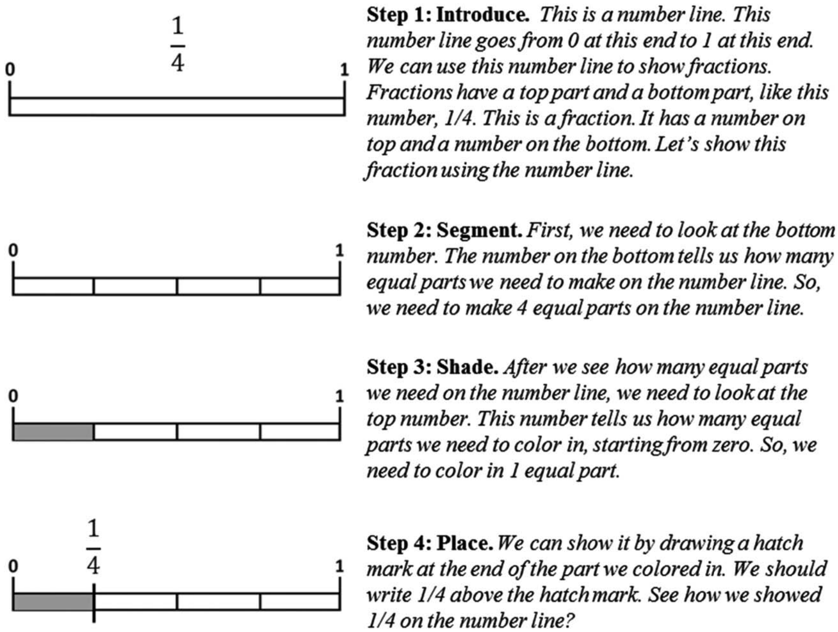

Figure 1. Number line training procedure.

amount of the circle. The fractions estimated were the same as those in the number line task and PAE was calculated for each item (PAE = |estimated area percentage − correct percentage|). The online supplemental materials provide detailed scoring instructions. This measure showed good reliability ($\alpha_{pretest}$ = .78; $\alpha_{posttest}$ = .84).

Magnitude comparison. In the magnitude comparison task (e.g., Bailey et al., 2015), children were shown two fractions and asked to point to the larger fraction (see Figure 3; see the online supplemental materials for all items). This is a valid measure, widely used by others to gauge children's ability to compare fraction magnitudes, with performance varying with the individu-

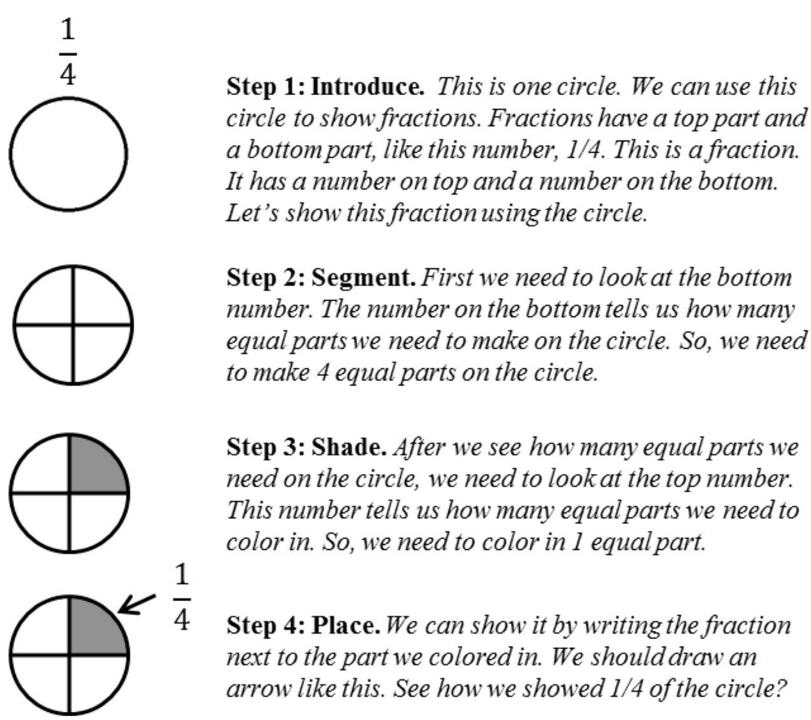

Figure 2. Area model training procedure.

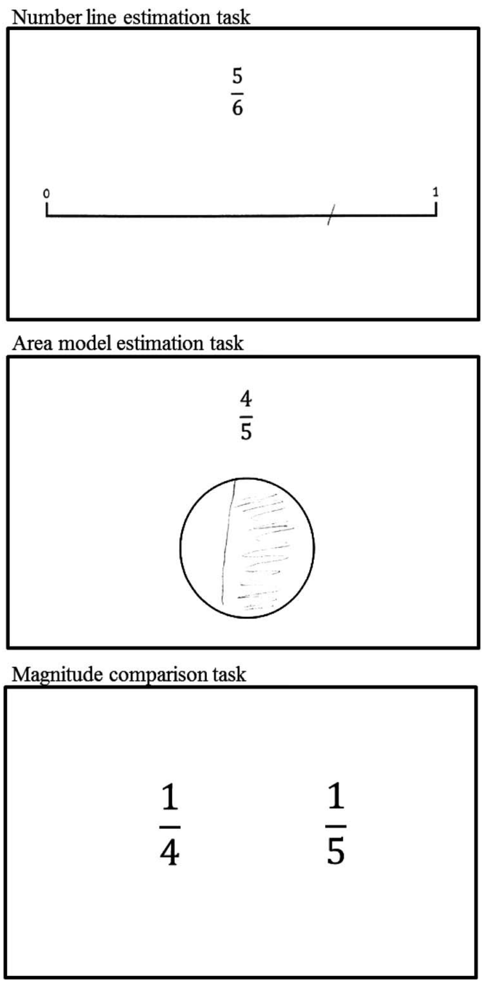

Figure 3. Fraction pre- and posttest measure examples.

al's experience with fractions and the size of the fractions (e.g., Siegler & Pyke, 2013). Pre- and posttest magnitude comparison performance was significantly correlated, $r(112) = 0.65$, $p < .001$. There were 20 items, each presented on a separate sheet of paper. Fraction denominators ranged from 2 to 7. The side of the correct response (left vs. right) was counterbalanced.

A common strategy used to complete this task is a bigger-whole-number strategy (e.g., Ni & Zhou, 2005), involving comparing the whole numbers of the numerators and denominators of each fraction to decide which fraction is larger (e.g., Clarke & Roche, 2009). For most items on our task (10 "inconsistent" items; e.g., 1/3 vs. 1/6), this strategy would yield the incorrect response. For some items (six "consistent" items; e.g., 6/7 vs. 2/6), this strategy would yield the correct response. In a few cases (four "ambiguous" items; e.g., 2/4 vs. 1/5), the usefulness of this strategy was ambiguous, in which comparing the numerators yields a different response than comparing the denominators. Reliability was good for each item type: inconsistent items, $\alpha_{pretest} = .97$, $\alpha_{posttest} = .96$; consistent items, $\alpha_{pretest} = .90$, $\alpha_{posttest} = .90$; ambiguous items, $\alpha_{pretest} = .84$, $\alpha_{posttest} = .84$.

Reading achievement. As a control, students completed the Letter-Word Identification subtest of the nationally normed Woodcock-Johnson Tests of Achievement, Fourth Edition (Schrank, Mather, & McGrew, 2014). This subtest requires students to read single letters and words of increasing difficulty.

Results

Preliminary Analyses

Means and correlations among all measures are displayed in Table 1 (for means by condition, see Table 2; for correlations within condition, see Table B of the online supplemental materials). Each fraction task was significantly correlated with the same task from pretest to posttest. However, none of the fraction tasks were significantly intercorrelated within pretest or posttest. Although previous work (e.g., Siegler & Pyke, 2013) observed significant, moderate correlations between number line estimation and magnitude comparison, we did not observe this, likely because children with minimal fraction magnitude knowledge may be using different task-specific strategies to solve these problems.

Main Analyses

We conducted analyses of covariance (ANCOVAs) for each of our posttest fraction magnitude tasks: number line estimation, area model estimation, and magnitude comparison. All ANCOVAs included condition as a between-subjects factor, and pretest performance on the same task, reading achievement, and child age as covariates. We included age, rather than grade, as a covariate because we expected age to be more sensitive to developmental differences than grade. Replacing age with grade did not change the results.[3]

Number line estimation. We predicted that the number line training condition would outperform the other two conditions on number line estimation. Posttest number line PAE significantly differed by condition, $F(2, 102) = 6.41$, $p = .002$, $\eta_p^2 = .11$ (Figure 4, Panel A). At posttest, the number line training group performed significantly better on number line estimation (i.e., had lower PAEs; adjusted $M = 0.24$) than those in the crossword control (adjusted $M = 0.32$, $p < .001$, $d = 0.84$) and those in the area model training (adjusted $M = 0.29$, $p = .025$, $d = 0.54$). The area model and crossword conditions did not significantly differ on posttest number line estimation ($p = .226$, $d = 0.29$).

To further interpret the effect size in terms of pre- to posttest change, we compared the average pretest number line PAE across conditions ($M = 0.29$) with the adjusted posttest score for each condition. The number line training group had an average posttest score (adjusted $M = 0.24$) that was 4.92 percentage points lower (i.e., better) than the average pretest score. Area model training and crossword control conditions

[3] Specifically, for our main outcome variables (number line PAE, area model PAE, and magnitude comparison), all effects reported to be statistically significant ($p < .05$) remained so when replacing age with grade level. When controlling for grade instead of age, there was also a significant difference in posttest area model estimation between children in the number line training and those in the area model training ($p = .049$), whereas when controlling for child age, this difference was marginally significant ($p = .060$).

Table 1
Means (SDs) and Correlations Among All Pretest and Posttest Measures

Measure	Mean (*SD*)	*N*	1	2	3	4	5	6	7	8	9
1. Child gender (female = 0, male = 1)	.39 (.49)	114	—								
2. Child grade level (second or third)	2.45 (.50)	114	−.01	—							
3. Child years of age	8.55 (.58)	114	.00	.37***	—						
Pretest measures											
4. Number line estimation (PAE)	.29 (.07)	112	−.05	.15	−.06	—					
5. Area model estimation (PAE)	.22 (.10)	114	−.03	−.20*	−.13	.05	—				
6. Magnitude comparison accuracy	.41 (.12)	114	−.08	.07	.11	−.10	.13	—			
7. Reading achievement (W score)	480.6 (24.8)	109	−.10	.08	.32***	−.15	−.21*	.20*	—		
Posttest measures											
8. Number line estimation (PAE)	.28 (.10)	113	.10	.18	−.01	.33***	.19*	−.08	−.18	—	
9. Area model estimation (PAE)	.18 (.10)	113	.12	−.05	.04	.15	.37***	.08	−.34***	.10	—
10. Magnitude comparison accuracy	.42 (.11)	114	−.00	.06	.07	−.25**	−.07	.65***	.17	−.17	−.15

Note. PAE = percent absolute error.
* $p < .05$. ** $p < .01$. *** $p < .001$.

had higher (i.e., worse) PAEs at posttest than at pretest, by .04% and 2.78%, respectively.[4]

We also examined performance on estimating unit fractions and nonunit fraction locations. Unit fractions are the building blocks for all other fractions, and the ability to estimate unit fractions on a number line is particularly predictive of algebra readiness (Booth & Newton, 2012). The number line training condition outperformed the others at posttest for unit fractions, $F(2, 102) = 7.22$, $p = .001$, $\eta_p^2 = .12$. There was no significant condition difference on nonunit fractions, $F(2, 102) = 1.80$, $p = .170$, $\eta_p^2 = .03$.

Area model estimation. We predicted that the area model training group would outperform the other groups on area model estimation. Children's posttest area model PAE differed by condition, $F(2, 102) = 5.11$, $p = .008$, $\eta_p^2 = .09$ (Figure 4, Panel B). At posttest, the area model training group performed significantly better at estimating fractions on the area model (adjusted $M = 0.14$) than the crossword control (adjusted $M = 0.21$, $p = .002$, $d = 0.77$) and marginally better than the number line training group (adjusted $M = 0.18$, $p = .060$, $d = 0.44$). The number line training and crossword conditions did not significantly differ on posttest area model estimation ($p = .184$, $d = 0.31$).

We compared the average pretest area model PAE across conditions ($M = 0.22$) with the adjusted posttest score for each condition. The area model training condition posttest score (adjusted $M = 0.14$) was 7.93 percentage points lower (i.e., better) than the average pretest. The number line and crossword conditions had smaller improvements in PAE from pretest to posttest, by 3.82% and 1.02%, respectively.[5]

There was no significant condition effect on estimating area model unit fractions at posttest, $F(2, 102) = 1.23$, $p = .297$, $\eta_p^2 = .02$. There was a significant condition effect on estimating nonunit fractions, $F(2, 102) = 6.11$, $p = .003$, $\eta_p^2 = .11$, with the area model condition performing significantly better than the crossword condition ($p = .001$) and marginally better than the number line condition ($p = .088$); the number line condition performed marginally better than the crossword condition ($p = .072$).

Magnitude comparison. We predicted that only the number line training group would transfer their knowledge to comparing fraction magnitudes, an untrained task. This key prediction was confirmed: children's performance on the posttest fraction magnitude comparison task differed by condition, $F(2, 103) = 5.10$, $p = .008$, $\eta_p^2 = .09$ (Figure 4, Panel C). At posttest, the number line group performed significantly better at comparing fractions (adjusted $M = 44.8\%$) compared with the area model (adjusted $M = 38.7\%$, $p = .002$, $d = 0.75$) and crossword conditions (adjusted $M = 40.9\%$, $p = .045$, $d = 0.48$). The area model and crossword conditions did not significantly differ on posttest magnitude comparison ($p = .273$, $d = 0.27$).

The number line training group had an average posttest score (adjusted $M = 44.8\%$) that was 3.98 percentage points higher (i.e., better) than the average pretest score ($M = 40.8\%$).[6] However, the crossword control improved by only 0.08%. The area model training group had lower magnitude comparison scores at posttest than at pretest, by 2.10%.[7]

We examined the magnitude comparison task in more detail according to whether the "bigger whole number strategy" was consistent with the correct answer (six items), inconsistent (10 items), or ambiguous (four items). Because performance on each item type was non-normally distributed, we used binomial regressions to model condition effects, controlling for child age, reading achievement, and pretest score on the same items.

[4] When examining unadjusted scores (see Table 2), children who received number line training significantly improved in their number line estimation from pretest to posttest, $t(38) = 2.59$, $p = .013$, $d = .42$, whereas those in the other training groups did not ($ps > .05$).

[5] In comparing unadjusted pretest to posttest performance (see Table 2), only children who received the area model training improved in their area model estimation, $t(35) = 4.16$, $p < .001$, $d = .70$.

[6] The pretest mean differs slightly from Table 1 because it includes only the $n = 108$ subjects included in this analysis.

[7] When comparing unadjusted scores (see Table 2), children in the number line condition performed marginally better on magnitude comparison at posttest than at pretest, $t(38) = 1.76$, $p = .086$, $d = .28$. In the area model and crossword puzzle groups, change from pretest to posttest was not significantly different from zero ($ps > .60$).

Table 2
Means and SDs for All Measures by Condition

Measure	Number line condition (n = 39) M (SD)	Area model condition (n = 37) M (SD)	Crossword condition (n = 38) M (SD)	Test of condition difference
Demographics				
Child gender	24 female; 15 male	22 female; 15 male	23 female; 15 male	$\chi^2(2) = .03$, $p = .983$
Child grade level	22 second; 17 third	18 second; 19 third	23 second; 15 third	$\chi^2(2) = 1.10$, $p = .577$
Child years of age	8.58 (.66)	8.50 (.60)	8.58 (.49)	$F(2, 111) = .219$, $p = .803$
Pretest measures				
Number line estimation (PAE)	.28 (.07)	.28 (.08)	.29 (.07)	$F(2, 109) = .271$, $p = .763$
Unit PAE	.32 (.14)	.32 (.17)	.31 (.14)	$F(2, 109) = .043$, $p = .958$
Non-Unit PAE	.26 (.10)	.26 (.08)	.28 (.10)	$F(2, 109) = .891$, $p = .413$
Area model estimation (PAE)	.19 (.11)	.24 (.11)	.22 (.09)	$F(2, 111) = 2.03$, $p = .136$
Unit PAE	.19 (.19)	.25 (.18)	.20 (.15)	$F(2, 111) = 1.11$, $p = .335$
Non-Unit PAE	.19 (.13)	.23 (.11)	.23 (.11)	$F(2, 111) = 1.71$, $p = .185$
Magnitude comparison overall PA	.43 (.13)	.38 (.12)	.43 (.12)	$F(2, 111) = 1.78$, $p = .173$
Consistent PA	.83 (.32)	.93 (.19)	.85 (.29)	$F(2, 111) = 1.51$, $p = .225$
Inconsistent PA	.17 (.34)[a]	.01 (.03)[b]	.09 (.25)[a,b]	$F(2, 111) = 4.08$, $p = .019$
Ambiguous PA	.47 (.40)	.49 (.41)	.63 (.41)	$F(2, 111) = 1.76$, $p = .178$
Reading achievement (W score)	484.21 (22.71)	478.14 (23.72)	479.03 (28.14)	$F(2, 106) = .650$, $p = .524$
Posttest measures				
Number line estimation (PAE)	.24 (.12)[a]	.29 (.09)[b]	.32 (.08)[b]	$F(2, 110) = 7.73$, $p = .001$
Unit PAE	.26 (.17)[a]	.36 (.14)[b]	.38 (.12)[b]	$F(2, 110) = 8.29$, $p < .001$
Non-Unit PAE	.22 (.11)	.23 (.10)	.28 (.12)	$F(2, 110) = 2.88$, $p = .060$
Area model estimation (PAE)	.17 (.11)[a,b]	.15 (.10)[a]	.21 (.09)[b]	$F(2, 110) = 3.57$, $p = .032$
Unit PAE	.17 (.16)	.14 (.15)	.18 (.12)	$F(2, 110) = .76$, $p = .469$
Non-Unit PAE	.16 (.12)[a]	.16 (.09)[a]	.23 (.14)[b]	$F(2, 110) = 4.71$, $p = .011$
Magnitude comparison overall PA	.46 (.12)[a]	.38 (.09)[b]	.42 (.10)[a,b]	$F(2, 111) = 6.02$, $p = .003$
Consistent PA	.83 (.35)	.95 (.11)	.88 (.25)	$F(2, 111) = 1.80$, $p = .169$
Inconsistent PA	.19 (.35)[a]	.01 (.04)[b]	.10 (.26)[a,b]	$F(2, 111) = 4.33$, $p = .015$
Ambiguous PA	.58 (.41)	.43 (.39)	.53 (.41)	$F(2, 111) = 1.35$, $p = .262$

Note. For rows with a significant overall F test, conditions that significantly differ ($p < .05$) are labeled with different letters; conditions showing no significant difference ($p > .05$) are labeled with the same letter. PAE = percent absolute error; PA = percent accuracy.

On consistent items, children's performance was near ceiling at pretest ($M = 0.86$, $SD = .28$) and posttest ($M = 0.88$, $SD = .27$). Posttest scores did not significantly differ between number line and area model trainings, B = .75, *Wald* (1, $N = 109$) = 3.48, $p = .062$, or between number line and crossword conditions, B = .38, *Wald* (1, $N = 109$) = 1.23, $p = .268$.

On inconsistent items, children's performance was near floor at pretest ($M = 0.09$, $SD = .26$) and posttest ($M = 0.11$, $SD = .27$). The number line training group performed significantly better at posttest than the area model training group, B = 1.40, *Wald* (1, $N = 109$) = 6.34, $p = .012$. The other conditions did not significantly differ ($ps > .05$). Although all conditions had posttest adjusted means that were lower than the average pretest score ($M = 9.36\%$), the number line training had the highest adjusted posttest (adjusted $M = 7.03\%$), followed by the crossword condition (adjusted $M = 5.16\%$), then the area model condition (adjusted $M = 1.83\%$).

On ambiguous items, scores were intermediate at pretest ($M = 0.51$, $SD = 0.41$) and posttest ($M = 0.50$, $SD = 0.41$). There was a significant effect of condition on posttest scores (see Figure 5). The number line condition performed significantly better on the ambiguous comparison items than the area model training, B = 0.99, *Wald* (1, $N = 109$) = 11.12, $p = .001$, and the Crossword control, B = 1.05, *Wald* (1, $N = 109$) = 12.15, $p < .001$. The area model and crossword conditions did not significantly differ, B = −.06, *Wald* (1, $N = 109$) = .04, $p = .837$. The number line training group's ambiguous comparison posttest scores (adjusted $M = 65.1\%$) were, on average, 13.92 percentage points higher (i.e., better) than the average pretest ($M = 51.1\%$). In contrast, the area model and crossword conditions had worse ambiguous comparison scores at posttest than at pretest, by 10.16% and 11.65%, respectively.

Discussion

We show, for the first time, that the number line has a causal influence on fraction magnitude knowledge in young children. We compared number line training with a well-matched area model training to isolate each representation's causal role in children's fraction magnitude concepts. Both groups improved on the representation that they were taught—compared with the control condition, number line training improved children's number line estimation and area model training improved children's area model estimation. However, only the number line training group transferred their learning to fraction magnitude comparisons, an untrained task.

Although young children often incorrectly apply whole-number knowledge to fractions (e.g., Ni & Zhou, 2005), number line training combated against this bias. On number line estimation, unit fractions are most likely to be impacted by whole-number bias (e.g., children mistakenly believe that [1/7] is large, whereas [1/3] is small). Number line training led to higher performance at estimating unit fractions on the number line, a skill strongly related to algebra readiness (Booth & Newton, 2012). In contrast, area

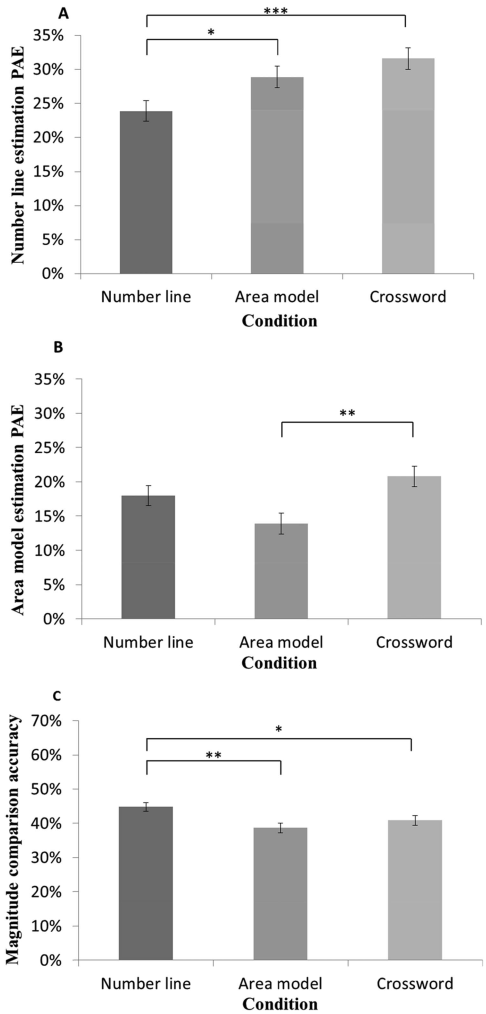

Figure 4. Posttest fraction number line estimation (Panel A), area model estimation (Panel B), and magnitude comparison (Panel C) by condition. Scores control for pretest on the same measure, child age, and reading achievement. * $p < .05$. ** $p < .01$. *** $p < .001$.

model training did not improve performance at estimating unit fractions on the area model. On the magnitude comparison task, number line training led to higher performance on items inconsistent with the whole-number bias (e.g., [1/4] vs. [1/5]), although performance remained low overall. Finally, number line training led to the most dramatic improvement on magnitude comparison when whole-number strategies were unavailable (ambiguous items). These results suggest that number line training reduces children's susceptibility to whole-number interference, perhaps by drawing on preexisting horizontal space and magnitude associations.

We note that, on ambiguous magnitude comparison items, the area model and crossword conditions performed worse at posttest than pretest. This finding is likely because of greater fatigue or lack of attention during the posttest than the pretest, perhaps resulting from the sustained effort and attention required during the training, or from the cumulative effects of participation across two sessions. Nevertheless, the conceptual benefits of number line training appeared to counteract the fatigue effects to produce improvement; in the absence of fatigue effects, improvement from pretest to posttest in the number line training would likely be even larger.

Teaching children how to represent fractions on a number line improves magnitude concepts. This study is the first to demonstrate that even children with minimal exposure to fractions can benefit from number line training. Several properties of the number line make it more optimal for teaching fraction magnitudes than the area model. Its linear and unidimensional nature, and the link it provides between spatial extents and number magnitudes, might render it an intuitive representation for relational fraction magnitude inferences. The number line presents children with a physical representation that matches the mental representation that promotes math achievement (Ramani & Siegler, 2008).

This study opens up new directions for theory and practice. Given that connecting multiple visual representations has been shown to enhance fraction learning (Cramer & Wyberg, 2009), the effectiveness of the number line we used in training (a thin rectangle) might be the result of combining number line elements (linearity and endpoints) with area model elements (two dimensions) in a single representation. Future studies that train children using a unidimensional number line can help determine whether the number line used for training in this particular study (which included two dimensions) contributed to our findings, or whether the number line is more generally beneficial for children's fraction magnitude concepts. Another interesting possibility is that exposure to the number line and area models at pretest allowed children to benefit from the slightly two-dimensional number line used during training. A future study that eliminates these pretests could help to determine whether prior exposure to both models was necessary to allow children to learn from the number line training.

More generally, further research can investigate what properties of the number line (i.e., labeled endpoints, left-to-right orientation, unidimensionality) make it fundamental for children's fraction magnitudes. Future studies can also ask whether the causal impact of the number line is limited to fraction magnitude concepts, or

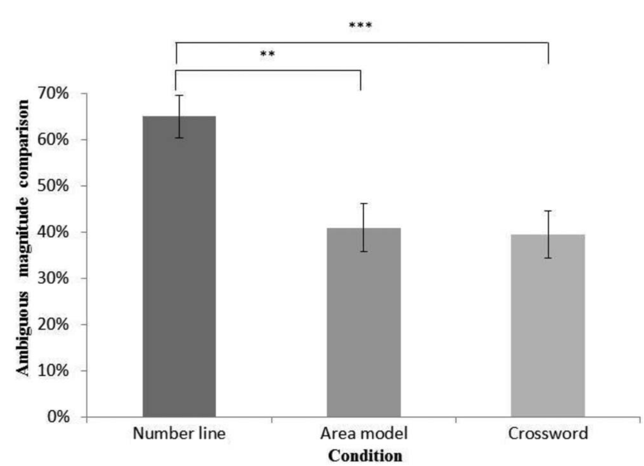

Figure 5. Posttest fraction magnitude comparison accuracy for ambiguous whole number strategy items, controlling for pretest performance on the same items, child age, and reading achievement. ** $p < .01$. *** $p < .001$.

whether it extends to fraction procedures, and can examine use of number line visualization strategies during fraction problem solving. From a practical perspective, we used a very brief training that likely underestimates the number line's potential to improve children's fraction magnitude concepts. Although we show that young children can learn from brief fraction training, performance was still poor at posttest, likely because of children's lack of fraction experience. Future research focused on practical applications can intervene with older children, increase the training dosage, and examine performance on delayed posttests to assess the number line's efficacy for enhancing children's fraction magnitude concepts over time. Future studies can also include other untrained tasks to assess the benefits of each training type for different fraction concepts and procedures.

In conclusion, the number line, in bridging numerical and spatial properties, is crucial for children's fraction magnitude concepts, facilitating deeper conceptual understanding that fractions represent magnitudes that can be ordered and compared.

References

Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology, 113,* 447–455. http://dx.doi.org/10.1016/j.jecp.2012.06.004

Bailey, D. H., Zhou, X., Zhang, Y., Cui, J., Fuchs, L. S., Jordan, N. C., . . . Siegler, R. S. (2015). Development of fraction concepts and procedures in U.S. and Chinese children. *Journal of Experimental Child Psychology, 129,* 68–83. http://dx.doi.org/10.1016/j.jecp.2014.08.006

Booth, J. L., & Newton, K. J. (2012). Fractions: Could they really be the gatekeeper's doorman? *Contemporary Educational Psychology, 37,* 247–253. http://dx.doi.org/10.1016/j.cedpsych.2012.07.001

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development, 79,* 1016–1031. http://dx.doi.org/10.1111/j.1467-8624.2008.01173.x

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

Braithwaite, D. W., & Goldstone, R. L. (2013). Integrating formal and grounded representations in combinatorics learning. *Journal of Educational Psychology, 105,* 666–682. http://dx.doi.org/10.1037/a0032095

Clarke, D. M., & Roche, A. (2009). Students' fraction comparison strategies as a window into robust understanding and possible pointers for instruction. *Educational Studies in Mathematics, 72,* 127–138. http://dx.doi.org/10.1007/s10649-009-9198-9

Common Core State Standards Initiative. (2015). *Preparing America's students for college & career: Mathematics standards*. Retrieved from http://www.corestandards.org/Math/

Cramer, K., & Wyberg, T. (2009). Efficacy of different concrete models for teaching the part-whole construct for fractions. *Mathematical Thinking and Learning, 11,* 226–257. http://dx.doi.org/10.1080/10986060903246479

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122,* 371–396. http://dx.doi.org/10.1037/0096-3445.122.3.371

de Hevia, M. D., & Spelke, E. S. (2010). Number-space mapping in human infants. *Psychological Science, 21,* 653–660. http://dx.doi.org/10.1177/0956797610366091

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology, 123,* 53–72. http://dx.doi.org/10.1016/j.jecp.2014.01.013

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. http://dx.doi.org/10.1037/a0032446

Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008). Development of number line representations in children with mathematical learning disability. *Developmental Neuropsychology, 33,* 277–299. http://dx.doi.org/10.1080/87565640801982361

Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2012). The relation between spatial skill and early number knowledge: The role of the linear number line. *Developmental Psychology, 48,* 1229–1241. http://dx.doi.org/10.1037/a0027433

Hoffmann, D., Hornung, C., Martin, R., & Schiltz, C. (2013). Developing number-space associations: SNARC effects using a color discrimination task in 5-year-olds. *Journal of Experimental Child Psychology, 116,* 775–791. http://dx.doi.org/10.1016/j.jecp.2013.07.013

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103,* 17–29. http://dx.doi.org/10.1016/j.jecp.2008.04.001

Institute of Education Sciences. (2010). *Developing effective fraction instruction for kindergarten through 8th grade*. Retrieved from https://ies.ed.gov/ncee/wwc/pdf/practice_guides/fractions_pg_093010.pdf

LeFevre, J.-A., Jimenez Lira, C., Sowinski, C., Cankaya, O., Kamawar, D., & Skwarchuk, S.-L. (2013). Charting the role of the number line in mathematical development. *Frontiers in Psychology, 4,* 641. http://dx.doi.org/10.3389/fpsyg.2013.00641

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science, 21,* 873–881. http://dx.doi.org/10.1177/0956797610370158

National Council of Teachers of Mathematics. (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics*. Washington, DC: Author.

Newcombe, N. S., Levine, S. C., & Mix, K. S. (2015). Thinking about quantity: The intertwined development of spatial and numerical cognition. *WIREs Cognitive Science, 6,* 491–505. http://dx.doi.org/10.1002/wcs.1369

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

Opfer, J. A., & Siegler, R. S. (2012). Development of quantitative thinking. In K. J. Holyoak & R. G. Morrison (Eds.), *Oxford handbook of thinking and reasoning* (pp. 585–605). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199734689.013.0030

Patro, K., & Haman, M. (2012). The spatial-numerical congruity effect in preschoolers. *Journal of Experimental Child Psychology, 111,* 534–542. http://dx.doi.org/10.1016/j.jecp.2011.09.006

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79,* 375–394. http://dx.doi.org/10.1111/j.1467-8624.2007.01131.x

Schneider, M., Grabner, R. H., & Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: Their interrelations in grades 5 and 6. *Journal of Educational Psychology, 101,* 359–372. http://dx.doi.org/10.1037/a0013840

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV Tests of Achievement*. Rolling Meadows, IL: Riverside.

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

Siegler, R. S., & Lortie-Forgues, H. (2014). An integrative theory of numerical development. *Child Development Perspectives, 8,* 144–150. http://dx.doi.org/10.1111/cdep.12077

Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding of fractions. *Developmental Psychology, 49,* 1994–2004. http://dx.doi.org/10.1037/a0031200

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62,* 273–296. http://dx.doi.org/10.1016/j.cogpsych.2011.03.001

Solomon, T. L., Vasilyeva, M., Huttenlocher, J., & Levine, S. C. (2015). Minding the gap: Children's difficulty conceptualizing spatial intervals as linear measurement units. *Developmental Psychology, 51,* 1564–1573. http://dx.doi.org/10.1037/a0039707

Torbeyns, J., Schneider, M., Xin, Z., & Siegler, R. S. (2015). Bridging the gap: Fraction understanding is central to mathematics achievement in students from three different continents. *Learning and Instruction, 37,* 5–13. http://dx.doi.org/10.1016/j.learninstruc.2014.03.002

Vamvakoussi, X., & Vosniadou, S. (2010). How many decimals are there between two fractions? Aspects of secondary school students' understanding of rational numbers and their notation. *Cognition and Instruction, 28,* 181–209. http://dx.doi.org/10.1080/07370001003676603

Received March 18, 2016
Revision received October 3, 2016
Accepted October 11, 2016 ■