



ELSEVIER

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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



Number line unidimensionality is a critical feature for promoting fraction magnitude concepts



Elizabeth A. Gunderson^{a,*}, Noora Hamdan^a, Lindsey Hildebrand^b, Victoria Bartek^a

^a Department of Psychology, Temple University, Philadelphia, PA 19122, USA

^b Department of Psychology, Boston College, Chestnut Hill, MA 02467, USA

ARTICLE INFO

Article history:

Received 22 August 2018

Revised 27 June 2019

Keywords:

Number line

Area model

Fractions

Mathematical development

Magnitude

Numerical Cognition

ABSTRACT

Children's ability to estimate fractions on a number line is strongly related to algebra and overall high school math achievement, and number line training leads to better fraction magnitude comparisons compared with area model training. Here, we asked whether unidimensionality is necessary for the number line to promote fraction magnitude concepts and whether left-to-right orientation and labeled endpoints are sufficient. We randomly assigned second- and third-graders ($N = 148$) to one of four 15-min one-on-one, experimenter-led trainings. Three number line trainings had identical scripts, where the experimenter taught children to segment and shade the number line along the horizontal dimension. The number line conditions varied only in the vertical dimension of the training number line: pure unidimensional number line (17.5 cm horizontal line), hybrid unidimensional number line (17.5 × 0.6 cm rectangle), and square number line (17.5 × 17.5 cm). In the area model condition, children were taught to segment and shade a square (17.5 × 17.5 cm) along both dimensions. The conditions significantly differed in posttest fraction magnitude comparison accuracy (a transfer task), controlling for pretest accuracy, reading achievement, and age. In preregistered analyses, the hybrid unidimensional number line condition significantly outperformed the square area model condition and the square number line condition. In exploratory analyses accounting for training protocol fidelity, these results held and the pure unidimensional number line also outperformed the area model condition on fraction magnitude comparisons. We argue that unidimensionality is a critical feature of the number line

* Corresponding author.

E-mail address: liz.gunderson@temple.edu (E.A. Gunderson).

for promoting fraction magnitude concepts because it aligns with a key concept—that real numbers, including fractions, can be ordered along a single dimension.

© 2019 Elsevier Inc. All rights reserved.

Introduction

Mathematical skills are strong predictors of college completion and adult earnings (Ritchie & Bates, 2013), and the U.S. lags behind many countries in students' mathematical achievement (Mullis, Martin, Foy, & Hooper, 2016). One major obstacle to students' success is failure to understand fraction magnitudes (Siegler, Fazio, Bailey, & Zhou, 2013). Although fractions are a major focus of instruction and show overall growth in fourth and fifth grades (Resnick et al., 2016), students show weak understanding of fractions even at the end of high school (Kloosterman, 2010). This is concerning given that fraction magnitude knowledge predicts algebra skills and high school math achievement (Booth, Newton, & Twiss-Garrity, 2014; Siegler et al., 2012). To improve these outcomes, understanding the cognitive skills underlying fraction magnitudes is crucial.

The ability to estimate fraction magnitudes on a number line is an important skill that indexes and improves fraction magnitude concepts (Hamdan & Gunderson, 2017b; Resnick et al., 2016; Siegler & Lortie-Forgues, 2014; Torbeyns, Schneider, Xin, & Siegler, 2015). Fraction number line estimation is correlated with fraction procedural skill (Bailey et al., 2015) and algebra learning (Booth et al., 2014). Furthermore, experimental training studies have established a causal relation between fraction number line estimation and fraction concepts (Fuchs et al., 2013; Hamdan & Gunderson, 2017b). In one study, at-risk fourth-graders who received a fraction curriculum that focused on number lines outperformed those in an *area model*-focused curriculum (Fuchs et al., 2013). In a brief controlled experiment, teaching second- and third-grade students to represent fractions on a number line led to significantly higher performance on a transfer task, fraction magnitude comparison, than teaching students to show fractions on a circle (Hamdan & Gunderson, 2017b). These results are notable because representing fractions as part of a shape (an area model) is the only incorporation of fractions in the U.S. Common Core State Standards in first and second grades, with number lines introduced only later (Center, 2010). These studies establish the benefits of an external number line representation for teaching fractions early and have proposed theoretical reasons to expect such benefits (Fuchs et al., 2013; Hamdan & Gunderson, 2017b; Siegler & Lortie-Forgues, 2014). The goal of the current study was to empirically test two theoretical explanations. Specifically, we tested whether unidimensionality is *necessary* and whether left-to-right orientation and labeled endpoints are *sufficient* for children to benefit from fraction number line training.

Unidimensionality

An external number line representation should benefit fraction magnitude concepts because the unidimensionality of the number line matches an important conceptual feature of all real numbers—that they can be ordered on a single dimension. A mature mental representation of whole-number magnitudes is well described by an analog mental number line, matching this conceptual feature of real numbers (Dehaene, 1992). Even preverbal infants associate line length—a unidimensional spatial property—with numerical magnitudes (Rugani & de Hevia, 2017). The association between numerical magnitude and linear space is refined and strengthened over childhood and acquires left-to-right directionality in cultures with left-to-right-oriented cultural practices (e.g., counting, reading) (McCrink & Opfer, 2014); we discuss the potential role of left-to-right orientation separately below. Furthermore, according to the integrative theory of numerical development, when students learn about new types of real numbers (e.g., fractions, decimals, negatives, irrationals), they progressively integrate each new type of number into a unified unidimensional mental number line (Siegler &

Lortie-Forgues, 2014). Evidence from educated adults supports this theory; when comparing fractions, adults' performance depended on the ratio of the numbers compared, consistent with an analog mental number line that includes fractions (Hurst & Cordes, 2016). Therefore, teaching children using an external unidimensional number line may be particularly beneficial because the external representation matches the desired mature internal mental representation.

Consistent with this, the number line was more effective at promoting fraction magnitude concepts than a two-dimensional circular area model (Hamdan & Gunderson, 2017b). However, because the number line and circle differ in several features, it is not clear whether unidimensionality was a necessary feature. Specifically, the number line involves left-to-right orientation and labeled endpoints, which may also be beneficial for fraction concepts. To test whether unidimensionality is necessary for promoting fraction concepts, we tested whether children would benefit from number line training without unidimensionality (Fig. 1C). If unidimensionality is a necessary feature of the number line, then a two-dimensional square number line should not be as effective as a unidimensional number line at promoting fraction magnitude concepts. In contrast, if unidimensionality is not necessary, and students gain information mainly from the number line's left-to-right orientation and labeled endpoints, then the two-dimensional square number line would be just as effective as a unidimensional number line.

We also examined whether the degree of unidimensionality would affect the number line's effectiveness. Prior research found that a slightly rectangular enclosed number line (referred to here as the hybrid unidimensional number line; Fig. 1B) was successful in promoting fraction magnitude concepts (Hamdan & Gunderson, 2017b). One reason why a hybrid unidimensional number line may be particularly helpful is that it maintains the property of unidimensionality and, when divided into segments using internal hash marks, directs the child's attention to horizontal space between hash marks as the meaningful unit. Directing attention to the spaces between hash marks may be important given that previous research on linear measurement showed that elementary school children often incorrectly focus on the hash marks themselves as objects to be counted (Solomon, Vasilyeva, Huttenlocher, & Levine, 2015). However, an alternative possibility is that the slight two-dimensionality of the hybrid unidimensional number line detracts from the conceptual alignment between the visual representation and the desired mental representation. If this were the case, we would expect a pure unidimensional number line (Fig. 1A) to be even more effective at promoting fraction magnitude concepts than the hybrid unidimensional number line. We tested these conflicting predictions by comparing the effects of pure and hybrid unidimensional number line trainings.

Left-to-right orientation and labeled endpoints

A second reason why the number line ought to benefit fraction magnitude concepts is that it draws on preexisting associations between numerical magnitudes and spatial locations (McCrink & Opfer, 2014; Rugani & de Hevia, 2017). Even infants and non-human animals associate smaller numerosities with left-hand space and larger numerosities with right-hand space (for a review, see Rugani & de Hevia, 2017). By toddlerhood, children who make this association have better number knowledge than children who do not associate left with smaller numbers and right with larger numbers (Opfer, Thompson, & Furlong, 2010; but note that this varies between cultures; Shaki, Fischer, & Göbel, 2012). Furthermore, when asked to estimate locations of whole numbers, kindergartners performed worse on a number line oriented from right to left than on a left-to-right number line (Ebersbach, 2015). This suggests that children and adults have strong expectations that smaller numbers appear on the left and larger numbers appear on the right, and number line training may be particularly impactful because it builds on these expectations. Supporting this, preschoolers taught about magnitudes of whole numbers using a left-to-right-oriented number board game improved in number line estimation and magnitude comparison more than those taught using a circular game (Siegler & Ramani, 2009).

Therefore, showing fractions using an external left-to-right-oriented number line aligns with children's expectations, and may aid children in making associations between fraction symbols and magnitudes. For example, when shown how to place fractions on a 0-to-1 number line, children may note that the fraction $1/6$ goes on the left side and the fraction $3/4$ goes on the right side, and the heuristic "left-smaller, right-larger" should benefit students in understanding fractions' relative magnitudes. In

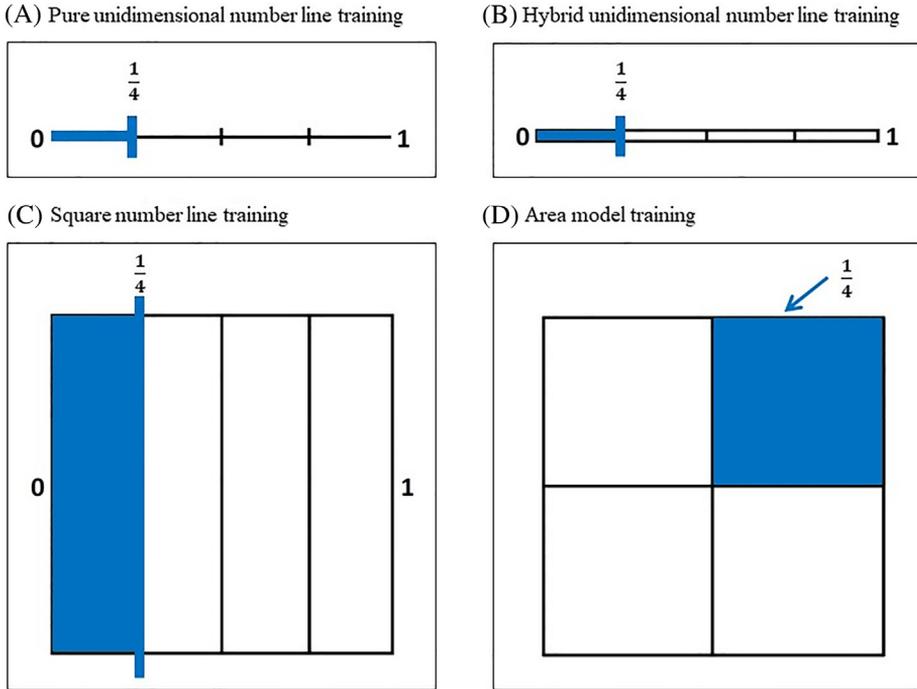


Fig. 1. Example of the representation used during training in each condition. (A) Pure unidimensional number line training. (B) Hybrid unidimensional number line training. (C) Square number line training. (D) Area model training.

contrast, the area model does not align with this association; it is equally correct to shade $1/6$ on the left or right side of a square. In the current study, we chose not to vary the orientation of the number line itself because prior research suggests that an intervention using a right-to-left-oriented number line would be harmful to learning (Ebersbach, 2015). Instead, we examined whether left-to-right orientation, in combination with labeled endpoints, is *sufficient* for children to benefit from the number line even when unidimensionality is not present (Fig. 1C). In our studies, to differentiate the number line trainings from area model training in terms of left-to-right orientation, we adopted the approach of Hamdan and Gunderson (2017b); in the area model training, the experimenter always shaded the area model in a clockwise manner, starting from the top right.

A third component of the external number line representation is that it provides numerical labels for specific benchmark locations, typically at the endpoints. When estimating whole numbers on a number line, children and adults use these endpoints (and sometimes additional benchmarks such as the midpoint) to estimate the proportional location on the line where a specific number should go (Barth & Paladino, 2011; Slusser & Barth, 2017; Slusser, Santiago, & Barth, 2013). In the case of fractions, most research has used external number lines with endpoints of 0–1, 0–2, or 0–5 (e.g., Booth et al., 2014; Hansen et al., 2015; Iuculano & Butterworth, 2011; Resnick et al., 2016). Theoretically, these endpoints should remind students that fractions lie on a continuum of real numbers (Siegler, Thompson, & Schneider, 2011). These endpoints provide points of comparison (e.g., $1/6$ is close to zero, $9/10$ is near 1) that may be quite helpful for understanding fraction magnitudes. In contrast, the area model does not provide this explicit support. When shading a square to represent fractions, it is implicit that a fully unshaded square represents 0 and a fully shaded square represents 1, but this may be difficult for children to intuit without explicit labels. Here, we asked whether the number line's labeled endpoints and left-to-right orientation together are *sufficient* for students to benefit from the external representation in the absence of unidimensionality. Notably, our study was *not* designed to test whether labeled endpoints and left-to-right orientation are *necessary* for the number line to promote fraction magnitude concepts; we return to this in the Discussion.

Specificity of the effects

In addition to examining the properties of the number line that affect children's fraction magnitude concepts, we also sought to characterize the change in children's fraction knowledge by including several different assessments. We asked children to show the magnitude of individual fractions by estimating their locations on a 0-to-1 number line and by shading a square. These tasks allowed us to determine whether children in each condition learned what they were taught.

We included both unit fractions (fractions with numerators of 1) and non-unit fractions (fractions with numerators other than 1). Research has suggested that understanding of unit fractions is a particularly strong predictor of algebra readiness (Booth & Newton, 2012). The whole-number bias may make it especially difficult for children to learn the magnitudes of unit fractions because the whole-number magnitude of the denominator is the inverse of the fraction's magnitude. Prior work showed an impact of number line training specifically on unit fractions (Hamdan & Gunderson, 2017b). If number line training reduces whole-number bias, then we would expect to see the same pattern here.

Importantly, we also included a fraction magnitude comparison task that was untrained in all conditions. Transfer to fraction magnitude comparisons would provide evidence that the training improved conceptual, and not just procedural, knowledge of fractions. On fraction magnitude comparison tasks, children frequently use component-based strategies, choosing the fraction with larger whole-number components (numerator and/or denominator) as being larger overall (e.g., Rinne, Ye, & Jordan, 2017). Therefore, we included fraction comparison items that varied in terms of whether choosing the fraction with the larger numerator and/or denominator would lead to the correct answer (consistent items), incorrect answer (inconsistent items), or neither (one fraction had the larger numerator and the other had the larger denominator; ambiguous items). Because young children are typically at ceiling on consistent items (Hamdan & Gunderson, 2017b), we did not expect the conditions to differ on these items. However, we expected number line training (more than area model training) to improve children's holistic concept of fractions' magnitudes. The ability to correctly compare fractions where the numerator and denominator differ in opposite directions is an important indicator of holistic representations of fractions (Rinne et al., 2017). Therefore, we expected to see improvement on these ambiguous items (Hamdan & Gunderson, 2017b). In addition, training using number lines could help children to overcome their bias to choose fractions with larger whole number components, leading to improvement on inconsistent items (Hamdan & Gunderson, 2017b).

The current study

This study was designed to isolate the features of the external number line representation that contribute to its effectiveness in promoting fraction magnitude concepts. We also sought to replicate the result from Hamdan and Gunderson (2017b) that hybrid number line training led to greater transfer to fraction magnitude comparison than area model training. We tested second- and third-graders, an age when children have had little exposure to fraction concepts but have been shown to be influenced by brief number line training (Hamdan & Gunderson, 2017b). We implemented a pretest-training-posttest design, where pretest and posttest were identical and children were randomly assigned to one of four 15-min experimenter-led trainings. The four training conditions (Fig. 1A–D) included two that conceptually replicated Hamdan and Gunderson (2017b) (hybrid unidimensional number line training [Fig. 1B] and area model training [Fig. 1D]). Our key new training condition was the square number line training (Fig. 1C), which includes left-to-right orientation and labeled endpoints but is not unidimensional, allowing us to isolate these properties. If unidimensionality is necessary for the number line to be effective, then the square number line training should not lead to transfer to fraction magnitude comparisons. In contrast, if left-to-right orientation and labeled endpoints are sufficient (even without unidimensionality), then the square number line training should lead to transfer.

We added a pure unidimensional number line training condition (Fig. 1A) to test whether the slight two-dimensionality of the hybrid unidimensional number line training helped students to focus on spaces rather than hash marks or whether this slight two-dimensionality detracted from its conceptual content as a unidimensional representation. Because the pure and hybrid unidimensional number line trainings were quite similar, it is also possible that their effects would not differ.

This study was preregistered on Open Science Framework (<https://osf.io/asgze/>) (Hamdan & Gunderson, 2017a). Our preregistered hypotheses are below. Hypotheses 1 to 3 delineate the alternative hypotheses tested by our new training conditions, and Hypotheses 4 to 6 lay out our expected replication of the results of Hamdan and Gunderson (2017b).

Hypothesis 1

If unidimensionality alone affects the number line's effectiveness, then the pure unidimensional number line should result in the best fraction magnitude comparison performance, followed by the hybrid unidimensional number line, followed by the square number line and square area model (both of which should have equal effects).

Hypothesis 2

If left-to-right orientation and endpoint labeling is sufficient for the number line to be effective, then the three number line conditions (pure unidimensional, hybrid unidimensional, and square) should have similar effects and all three should lead to greater performance on fraction magnitude comparison than the square area model.

Hypothesis 3

If the pure unidimensional number line is confusing because children view hash marks as objects instead of viewing spaces as objects, then we should see the hybrid unidimensional number line result in better outcomes than the pure unidimensional number line.

Hypothesis 4

Compared with children trained in the area model, children trained in the hybrid unidimensional number line should perform better at posttest in fraction number line estimation and magnitude comparison. Compared with children in the hybrid unidimensional number line condition, children in the area model condition should perform better at posttest in area model estimation.

Hypothesis 5

On the fraction magnitude comparison outcome, we expected that the predicted effects of condition would be present for ambiguous items (where a bigger whole-number strategy leads to ambiguous solutions) and inconsistent items (where a bigger whole-number strategy is inconsistent with answering correctly) but not for consistent items (where a bigger whole-number strategy is consistent with answering correctly).

Hypothesis 6

The pure unidimensional and hybrid unidimensional number line trainings should be more beneficial for estimating unit fractions (fractions with 1 in the numerator) than non-unit fractions on the number line, whereas the area model training should not have this effect on the area model.

Method

Participants

Participants were second- and third-graders ($N = 148$; 77 second-graders; 74 girls; $M_{\text{age}} = 8.36$ - years, $SD = 0.60$; $n_{\text{age}} = 145$) and were recruited from four elementary schools in a large city in the eastern United States.¹ Participants were from 12 classrooms (6 in each grade), with an average of 12.3 participants per classroom (range = 4–24). Participants' parents reported their children's race/eth-

¹ Our preregistered target sample size was 160 based on a power analysis using G*Power (analysis of covariance with four groups and two covariates, power = .80, $f = .26$, $\alpha = .05$). We stopped data collection with 162 participants, 160 of whom had completed our key dependent variable, posttest fraction magnitude comparison. However, due to an error in record keeping, 14 of these participants were removed from the dataset subsequent to the study's completion, leaving a final sample of 148, 146 of whom had completed the posttest fraction magnitude comparison task. A sensitivity analysis with the same parameters as our original power analysis indicated that this sample size ($N = 146$) has power to detect a medium effect size ($f = .277$).

nicity as 39.3% White, 24.8% Black/African American, 18.8% Hispanic, 12.8% multiple races/ethnicities, 2.6% Asian/Asian American, and 1.7% other ($n_{\text{race/ethnicity}} = 117$). Families were diverse in terms of annual family income ($M = \$60,357$, $SD = \$36,266$, range = <\$15,000 to >\$100,000; $n_{\text{family income}} = 105$) and parents' maximum level of education ($M = 15.22$ years [16 years represents a 4-year college degree], $SD = 2.53$, range = 10 years [high school] to 18 years [graduate degree]; $n_{\text{parents' education}} = 117$). Missing data were present on some measures. Our analyses used listwise deletion and included participants who had all available data for each analysis; therefore, the sample size varies by analysis.

Procedure

The study involved two sessions with a pretest-training-posttest design. Each session was conducted one on one with an experimenter for about 30 min. The experimenters were trained graduate and undergraduate students and research assistants. Experimenters were not blinded to the study's hypotheses or training conditions.

In Session 1, participants completed a reading achievement control measure followed by the fraction pretest measures.² In Session 2, on average 5 days later ($M = 5.18$ days, $SD = 5.94$), children completed one of four fraction training conditions (each ~15 min long), randomly assigned within classrooms. After training, in the same session, participants completed the fraction posttest. The fraction pretest and posttest were identical, were the same across conditions, and included number line estimation, area model estimation, and fraction magnitude comparison.³ Participants were randomly assigned to one of four task orders, and in each order participants completed the tasks in Session 2 in the reverse order of Session 1.⁴

Training conditions

The four training conditions were modeled after prior work (Hamdan & Gunderson, 2017b). In all four conditions, the experimenter showed children how to represent fractions on a specific representation (see Fig. 1) and asked children to practice representing fractions, with feedback. Training was completed in a paper packet. The scripts (available in Appendix A) were identical for the number line conditions. The script for the area model training was closely matched to the number line trainings in terms of language, training steps, and fractions trained.

In all four training conditions, students learned to represent $1/2$, $1/4$, $2/4$, $3/4$, $1/6$, $2/6$, $3/6$, $4/6$, and $5/6$ (in that order) on either a number line or an area model. For each number, children first used a presegmented representation and then an unsegmented representation. Experimenters showed children how to *segment* the representation based on the denominator (the "bottom number"), *shade* the number of segments in the numerator (the "top number"), and *place* the fraction by writing the fraction and drawing a hash mark (number line training) or arrow (area model training) to indicate its location. On each trial, students were shown a correct representation of the fraction and given feedback on whether their response matched the correct representation. If children's representation was incorrect, students were instructed to correct their response on a new page. For unit fractions ($1/2$, $1/4$, and $1/6$), students' own practice was preceded by experimenter-led demonstrations, where experimenters showed children how to represent the fraction on the number line or area model by *segmenting*, *shading*, and *placing* the fractions. We chose to train children using fractions with even-numbered denominators (2, 4, and 6) to allow for easy segmentation of the area model along both dimensions. We chose small denominators to reduce the procedural demands and overall complexity of the tasks given children's limited fraction experience.

² Five participants completed the reading achievement measure in Session 2 due to factors including students' time constraints (e.g., needing to return early for lunch) and experimenter error (failure to bring the Letter-Word Identification stimuli to the school during Session 1).

³ In addition to our main dependent variables, we also collected several exploratory measures: fraction arithmetic and strategy use during each task (child self-report and experimenter report). However, we determined that the experimental procedure for collecting strategy use was not implemented consistently, and we did not analyze it further. Information on the fraction arithmetic measure is provided in Appendix C.

⁴ Due to experimenter error, some participants' task order was not recorded ($n = 22$).

The number line trainings differed from the area model training in several key ways. First, the number line was labeled as going from 0 (on the left) to 1 (on the right), whereas the area model was not. Second, in the number line trainings, the experimenter stated that the number line must be shaded starting from 0, whereas the experimenter did not instruct children to shade the area model in any particular order (the experimenter consistently demonstrated shading in a clockwise manner starting at the top right but did not instruct children on this or correct them if they shaded in a different order). Third, in the number line conditions the experimenter instructed children to place the fraction by drawing a hash mark at the right end of the shaded area and writing the fraction above the hash mark, whereas in the area model condition the experimenter instructed children to place the fraction by drawing an arrow pointing to the shaded area and writing the fraction next to the arrow. The experimenter modeled drawing a diagonal arrow (see Fig. 1D) but did not correct children if they drew a horizontal or vertical arrow to indicate the shaded part of the square.

Pure unidimensional number line training

In the pure unidimensional number line training ($n = 35$), the experimenter showed children how to segment, shade, and place fractions on a number line that was shown as a single horizontal line labeled 0 at the left and 1 at the right (Fig. 1A). This number line was 17.5 cm wide. In the segment step the hash marks extended above and below the line, and in the shade step the shading was visible above and below the line (approximately the same thickness as the hybrid unidimensional number line training).

Hybrid unidimensional number line training

In the hybrid unidimensional number line training ($n = 41$), the number line was a thin horizontal rectangle (17.5 cm wide \times 0.6 cm high; Fig. 1B) like the number line used in the fraction training study by Hamdan and Gunderson (2017b). It is referred to here as a “hybrid” because it contains properties of a unidimensional number line (minimal thickness and emphasis on horizontal dimension) with properties of the square number line (hash marks enclosed within the rectangle). In the segment step, the hash marks were drawn enclosed within the thin rectangle.

Square number line training

In the square number line training ($n = 36$), the number line was a square labeled 0 at the left side and 1 at the right side (Fig. 1C). The square was 17.5 cm wide and 17.5 cm high (the same dimensions as the square used in the area model training). The experimenter referred to it as a “number line” and treated it in the same way as the other two number line trainings.

Area model training

In the area model training ($n = 36$), children were taught how to segment and shade a square (17.5 cm wide \times 17.5 cm high). The experimenter always demonstrated segmenting the square on both vertical and horizontal dimensions and demonstrated shading the square clockwise, starting at the top right; children were also shown a correct response for each fraction that reflected this procedure. The experimenter did not correct children if they chose to segment or shade in an accurate but different manner.

Measures

Fraction pretest and posttest measures

All children completed the same items on the fraction pretest and posttest. The number line and area model estimation tasks were administered on paper and later scanned to PDF and measured by trained experimenters using Adobe Photoshop (see Appendix D for scoring details).

Number line estimation. To assess learning or near transfer, children estimated the locations of fractions on a number line. Each number line was a single horizontal line (17.5 cm long) labeled 0 on the left and 1 on the right (the same as the pure unidimensional number line training; Fig. 1A). The experimenter introduced the task by demonstrating how to place $1/2$ on the number line. There were 14 test items, including fractions that were trained (6 items with denominators 2, 4, and 6) and

untrained (8 items with denominators 3, 5, and 7).⁵ The fractions tested were 1/2, 1/3, 2/3, 1/4, 3/4, 1/5, 2/5, 4/5, 1/6, 4/6, 5/6, 1/7, 3/7, and 6/7. On each item, the fraction to be estimated was centered above the number line. Items were presented in a forward or backward pseudorandom order. Participants' response location was converted to a decimal (length from zero to hash mark divided by total length of the number line), and each item was scored based on the percentage absolute error (PAE): $|(\text{actual response} - \text{correct response})|/(\text{number line range})$. We used average PAE in our analyses. Reliability was adequate ($\alpha_{\text{pretest}} = .61$, $\alpha_{\text{posttest}} = .84$).

Area model estimation. Children also showed fractions by shading a square area model. The square was 17.5 cm wide and 17.5 cm high,⁶ and the fraction to be estimated was centered above the square (the same as the area model training; Fig. 1D). The experimenter demonstrated how to show 1/2 on the square (dividing the square with a vertical line and shading one segment). The 14 items were the same as in the number line estimation task. The area that children shaded was converted to a decimal (shaded area divided by total area of the square), and each item was scored using PAE: $|(\text{actual response} - \text{correct response})|/(\text{total area of square})$. We used average PAE in our analyses. Reliability was good ($\alpha_{\text{pretest}} = .90$, $\alpha_{\text{posttest}} = .90$).

Magnitude comparison. To assess transfer, children completed an untrained fraction magnitude comparison task (see Appendix B for all items). On each item, children were shown two fractions on a laptop and were asked to press a computer key to indicate which fraction was larger. The task included 28 items, and item order was fully randomized using E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002). The fractions all were smaller than 1 and included denominators 2, 3, 4, 5, 6, and 7. For 8 "consistent" items, the larger fraction also had the larger components when comparing numerators and comparing denominators. For 8 "inconsistent" items, the larger fraction had the smaller components. For 12 "ambiguous" items, one fraction had a larger numerator and the other had a larger denominator. We included more ambiguous items because prior work had found the greatest impact of fraction number line training on these items (Hamdan & Gunderson, 2017b). Reliability was good overall ($\alpha_{\text{pretest}} = .82$, $\alpha_{\text{posttest}} = .88$) and for each item type (consistent items: $\alpha_{\text{pretest}} = .82$, $\alpha_{\text{posttest}} = .84$; inconsistent items: $\alpha_{\text{pretest}} = .85$, $\alpha_{\text{posttest}} = .90$; ambiguous items: $\alpha_{\text{pretest}} = .88$, $\alpha_{\text{posttest}} = .91$).

Reading achievement

Reading achievement was included as a control measure to statistically equate the conditions on academic achievement in a non-numerical domain. Reading achievement was assessed using the Woodcock–Johnson IV Letter–Word Identification subtest (Schrank, Mather, & McGrew, 2014). In this nationally normed test, participants read individual words of increasing difficulty until they reach a basal level (6 sequential items correct at the beginning of the test) and a ceiling level (6 sequential items incorrect at the end of the test). Participants' raw scores were converted to *W* scores, Rasch-scaled scores appropriate for examining individual differences (Woodcock, 1999). Participants were excluded due to experimenter error in administering insufficient items to establish basal or ceiling criteria levels ($n = 21$), experimenter error in record keeping ($n = 20$), and other experimenter error ($n = 1$).

Data analysis plan

We first report descriptive statistics for all measures, by condition, and correlations between measures. We then report our preregistered analyses to test our main hypotheses (see preregistration at <https://osf.io/asgze/>). Finally, we present exploratory analyses, which were motivated by two factors. First, our preregistered control variable (reading achievement) was missing for an unexpectedly high

⁵ Although not part of our preregistered analysis plan, we explored whether the overall pattern of performance held within trained and untrained items on the number line estimation and area model estimation tasks. In both cases, the results reported in the main text (analysis of covariance results showing significant differences at posttest between conditions, controlling for reading achievement and age) were the same within both trained and untrained items.

⁶ In some cases, the square used for area model estimation was slightly larger (18.4 × 18.4 cm). In all cases, we calculated participants' area model estimation PAE using the actual size of the printed squares for those participants.

number of participants, leading to a lower than expected sample size after listwise deletion. We explored whether excluding this control measure (and, therefore, including more participants) would change the results. Second, our preregistration did not account for experimenter error during the training session. Therefore, we explored whether analyzing only those participants where the experimenter demonstrations adhered to the experimental protocol would yield the same pattern of results.

Results

Exclusion criteria

In some cases, a participant's data were missing on a task due to experimenter error, computer error, or the child's absence during the testing session. Based on our preregistered plan, we excluded data on a task when the participant did not complete more than half the items on that task. We also preregistered excluding participants' data if their reading achievement was 3 standard deviations below the mean; however, this was not the case for any participant.

Descriptive statistics

Descriptive statistics for all measures, by condition, as well as raw condition differences are reported in [Table 1](#). There were no significant differences between conditions at pretest, whereas there were significant effects of condition on each of our main dependent variables. Correlations between all measures (collapsed across conditions) are reported in [Table 2](#). Tests of raw pretest to posttest change, by condition, are reported in Appendix [Table C1](#).

Preregistered analyses

Following our preregistered analysis plan, we performed an analysis of covariance (ANCOVA) on each posttest measure, with condition as the independent variable, and pretest score on the same items, reading achievement, and age as covariates.⁷ We conducted pairwise comparisons between conditions based on the estimated marginal means from each ANCOVA. These pairwise comparisons follow our preregistered analysis plan unless otherwise noted in the text.

Number line estimation

Controlling for age, reading achievement, and pretest number line PAE, posttest number line estimation PAE significantly differed across conditions, $F(3, 89) = 7.71, p < .001, \eta_p^2 = .206$ ([Fig. 2A](#)). To follow up, we conducted pairwise comparisons of the estimated marginal means of each number line condition with the area model condition. These pairwise comparisons showed that the posttest number line estimation performance was significantly worse after area model training (adjusted $M = .27, SE = .02$) than pure unidimensional number line training (adjusted $M = .19, SE = .02; p = .011, d = 0.91$) and hybrid unidimensional number line training (adjusted $M = .18, SE = .02; p = .002, d = 1.09$) but did not significantly differ from the square number line training (adjusted $M = .27, SE = .02; p = 1.00, d = 0.02$).

We examined whether these effects were present for number line estimation of unit fractions (i.e., fractions with a numerator of 1; 6 items) and non-unit fractions (8 items). For number line estimation of unit fractions, an ANCOVA (controlling for age, pretest performance, and reading achievement) revealed a significant difference between conditions at posttest, $F(3, 90) = 6.50, p < .001, \eta_p^2 = .178$. For number line estimation of non-unit fractions, there was also a significant difference between conditions at posttest, $F(3, 85) = 4.40, p = .006, \eta_p^2 = .134$. The lack of a difference between unit and non-unit fractions was confirmed by a nonsignificant Fraction Type (unit or non-unit) \times Condition interaction on posttest number line estimation PAE, controlling for reading achievement, age, and pretest number line PAE, $F(3, 87) = 1.69, p = .175, \eta_p^2 = .055$.

⁷ Following our preregistered plan, we also reran our three main ANCOVAs (number line PAE, area model PAE, and magnitude comparison accuracy), replacing age with grade level, and the pattern of results was the same.

Table 1
Means (and standard deviations) for all measures by condition.

	Pure unidimensional NL training (n = 35)	Hybrid unidimensional NL training (n = 41)	Square NL training (n = 36)	Area model training (n = 36)	Test of condition difference
<i>Demographics</i>					
Child gender	19 female, 16 male	19 female, 22 male	19 female, 17 male	17 female, 19 male	$\chi^2(3) = 0.70, p = .873$
Child grade level	20 second-graders, 15 third-graders	19 second-graders, 22 third-graders	18 second-graders, 18 third-graders	20 second-graders, 16 third-graders	$\chi^2(3) = 1.14, p = .768$
Child years of age	8.34 (0.58)	8.35 (0.64)	8.41 (0.58)	8.34 (0.63)	$F(3, 141) = 0.11, p = .953, \eta_p^2 = .002$
<i>Pretest measures</i>					
Number line est. (PAE)	.28 (.09)	.29 (.09)	.26 (.09)	.27 (.05)	$F(3, 134) = 1.01, p = .389, \eta_p^2 = .022$
NL unit PAE	.30 (.14)	.31 (.13)	.27 (.13)	.33 (.13)	$F(3, 137) = 1.49, p = .219, \eta_p^2 = .032$
NL non-unit PAE	.27 (.12)	.26 (.11)	.25 (.11)	.23 (.07)	$F(3, 132) = 1.30, p = .278, \eta_p^2 = .029$
Area model est. (PAE)	.22 (.14)	.22 (.13)	.19 (.12)	.21 (.13)	$F(3, 133) = 0.41, p = .749, \eta_p^2 = .009$
AM unit PAE	.24 (.22)	.21 (.17)	.16 (.11)	.22 (.20)	$F(3, 133) = 1.26, p = .291, \eta_p^2 = .028$
AM non-unit PAE	.20 (.12)	.22 (.14)	.21 (.14)	.21 (.12)	$F(3, 133) = 0.15, p = .932, \eta_p^2 = .003$
Magnitude comparison acc.	.44 (.14)	.45 (.17)	.44 (.18)	.43 (.15)	$F(3, 138) = 0.10, p = .961, \eta_p^2 = .002$
MC consistent items acc.	.90 (.19)	.89 (.21)	.86 (.26)	.85 (.22)	$F(3, 138) = 0.47, p = .704, \eta_p^2 = .010$
MC inconsistent items acc.	.07 (.16)	.07 (.21)	.11 (.25)	.08 (.14)	$F(3, 138) = 0.28, p = .840, \eta_p^2 = .006$
MC ambiguous items acc.	.38 (.31)	.40 (.32)	.38 (.34)	.38 (.29)	$F(3, 138) = 0.04, p = .989, \eta_p^2 = .001$
Reading ach. (W score)	470.30 (26.09)	474.48 (29.66)	476.77 (28.93)	472.57 (26.07)	$F(3, 102) = 0.25, p = .865, \eta_p^2 = .007$
<i>Posttest measures</i>					
Number line est. (PAE)	.18 (.10) ^a	.17 (.10) ^a	.24 (.08) ^b	.27 (.08) ^b	$F(3, 138) = 9.83, p < .001, \eta_p^2 = .176$
NL unit PAE	.21 (.17) ^{a,b}	.16 (.12) ^a	.26 (.14) ^{b,c}	.30 (.14) ^c	$F(3, 139) = 7.27, p < .001, \eta_p^2 = .136$
NL non-unit PAE	.16 (.07) ^a	.17 (.12) ^a	.23 (.12) ^b	.24 (.10) ^b	$F(3, 136) = 4.46, p = .005, \eta_p^2 = .090$
Area model est. (PAE)	.18 (.11) ^a	.14 (.11) ^{a,b}	.12 (.09) ^b	.12 (.08) ^b	$F(3, 140) = 3.05, p = .031, \eta_p^2 = .061$
AM unit PAE	.19 (.16) ^a	.11 (.09) ^b	.11 (.12) ^b	.11 (.09) ^b	$F(3, 139) = 3.46, p = .018, \eta_p^2 = .069$
AM non-unit PAE	.18 (.09)	.16 (.14)	.12 (.09)	.14 (.09)	$F(3, 140) = 2.20, p = .091, \eta_p^2 = .045$
Magnitude comparison acc.	.49 (.18) ^a	.58 (.22) ^b	.47 (.19) ^a	.44 (.17) ^a	$F(3, 142) = 3.90, p = .010, \eta_p^2 = .076$
MC consistent items acc.	.84 (.22)	.92 (.14)	.84 (.28)	.87 (.25)	$F(3, 142) = 1.03, p = .384, \eta_p^2 = .021$
MC inconsistent items acc.	.13 (.23)	.19 (.33)	.14 (.25)	.08 (.22)	$F(3, 142) = 1.15, p = .333, \eta_p^2 = .024$
MC ambiguous items acc.	.51 (.31) ^{a,b}	.62 (.36) ^b	.44 (.35) ^a	.40 (.35) ^a	$F(3, 142) = 2.94, p = .035, \eta_p^2 = .059$

Note. For rows with a significant overall *F* test, conditions that significantly differ ($p < .05$) are labeled with different letters. NL, number line; AM, area model; MC, magnitude comparison; est., estimation; PAE, percentage absolute error; acc., accuracy; ach., achievement.

Table 2

Correlations among all measures.

	N	1	2	3	4	5	6	7	8	9
1. Child gender (female = 0, male = 1)	148	–								
2. Child grade level (second or third)	148	.01	–							
3. Child years of age	145	.11	.66***	–						
<i>Pretest measures</i>										
4. Number line estimation (PAE)	138	.00	–.06	–.06	–					
5. Area model estimation (PAE)	137	–.06	–.08	–.18*	.18*	–				
6. Magnitude comparison accuracy	142	.08	.02	–.08	–.12	.03	–			
7. Reading achievement (<i>W</i> score)	106	.00	.30**	.24*	–.15	–.27**	.14	–		
<i>Posttest measures</i>										
8. Number line estimation (PAE)	142	–.10	–.01	.03	.32***	.31***	–.20*	–.27**	–	
9. Area model estimation (PAE)	144	.04	–.25**	–.21*	.17	.53***	.09	–.32**	.19*	–
10. Magnitude comparison accuracy	146	.08	.06	.03	–.24**	–.09	.61***	.21*	–.39***	.03

Note. PAE, percentage absolute error.

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

Area model estimation

The conditions significantly differed on posttest area model estimation, $F(3, 92) = 6.79$, $p < .001$, $\eta_p^2 = .181$ (Fig. 2B). We contrasted the square area model with each of the number line conditions. Posttest area model estimation was better in the square area model condition (adjusted $M = .12$, $SE = .01$) than in the pure unidimensional number line condition (adjusted $M = .19$, $SE = .02$; $p < .001$, $d = 1.04$) but did not differ from the hybrid unidimensional number line condition (adjusted $M = .12$, $SE = .01$; $p = .908$, $d = 0.03$) or the square number line condition (adjusted $M = .11$, $SE = .02$; $p = .725$, $d = 0.10$). This pattern of performance was evident for both unit fractions, $F(3, 90) = 6.00$, $p < .001$, $\eta_p^2 = .167$, and non-unit fractions, $F(3, 92) = 4.55$, $p = .005$, $\eta_p^2 = .129$, in separate models. In a model with fraction type (unit or non-unit) as a factor, there was no significant interaction of Fraction Type \times Condition, $F(3, 91) = 2.63$, $p = .055$, $\eta_p^2 = .080$.

Magnitude comparison

Fraction magnitude comparison was our key dependent variable and a measure of transfer in all conditions. An ANCOVA on posttest fraction magnitude comparison accuracy, controlling for pretest accuracy, reading achievement, and age, revealed a significant effect of condition, $F(3, 93) = 3.56$, $p = .017$, $\eta_p^2 = .103$ (Fig. 2C). Following our preregistered plan, we conducted pairwise comparisons between the estimated marginal means of each condition. Participants in the hybrid unidimensional number line training condition (adjusted $M = .54$, $SE = .03$) performed significantly better than those in the area model training (adjusted $M = .43$, $SE = .03$; $p = .004$, $d = 0.84$), replicating prior work. The hybrid unidimensional number line training also outperformed the square number line training (adjusted $M = .44$, $SE = .03$; $p = .014$, $d = 0.75$). Performance did not significantly differ in the square number line training versus the area model training ($p = .749$, $d = 0.09$). In addition, performance on posttest magnitude comparison in the pure unidimensional number line condition (adjusted $M = .49$, $SE = .03$) did not significantly differ from the hybrid unidimensional number line condition ($p = .181$, $d = 0.39$), the square number line condition ($p = .238$, $d = 0.36$), or the square area model condition ($p = .113$, $d = 0.45$).

We also examined performance on each item type (consistent: the larger fraction has larger whole-number components; inconsistent: the larger fraction has smaller whole-number components; or ambiguous: comparing numerators and denominators would yield different answers). Because accuracy within item type was not normally distributed, we conducted binomial regressions, with condition as an independent variable, and covariates of pretest score on the same items, reading achievement, and age. There was no significant effect of condition for inconsistent items, $Wald(3, N = 100) = 3.90$, $p = .272$. There was a significant effect of condition for consistent items, $Wald(3,$

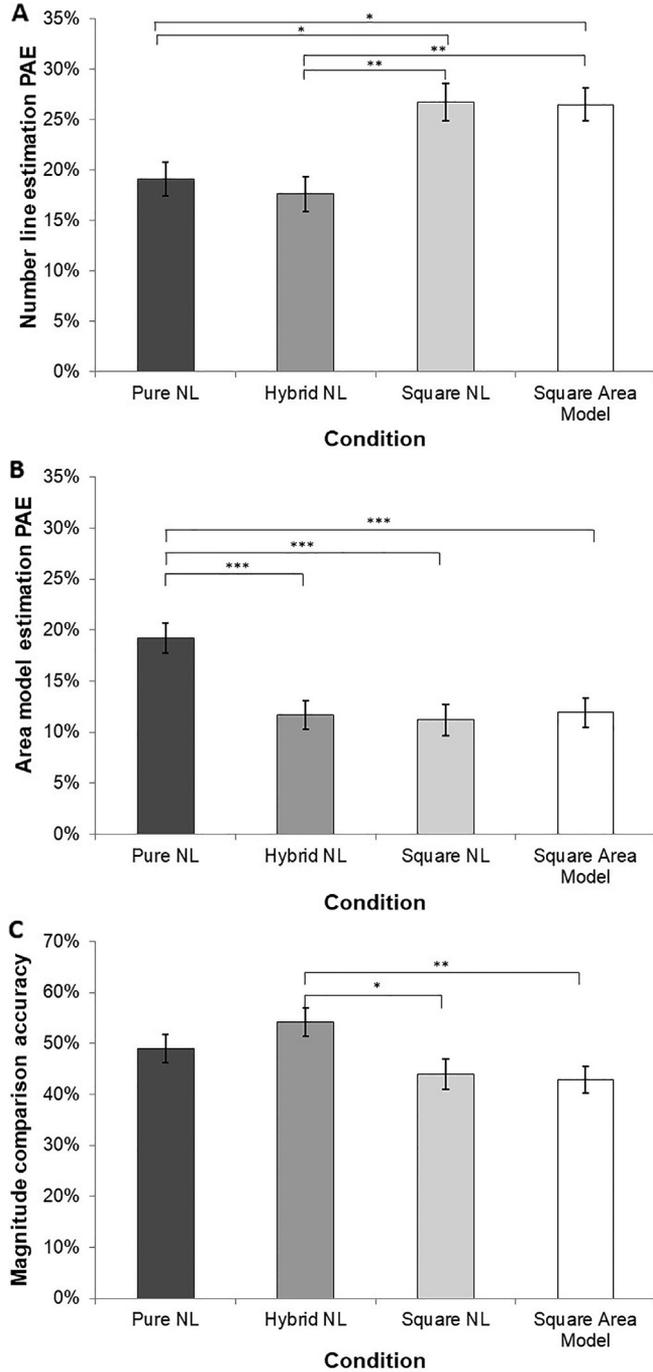


Fig. 2. Posttest fraction performance by training condition for number line (NL) estimation percentage absolute error (PAE) (A), area model estimation PAE (B), and magnitude comparison accuracy (C). Values are adjusted means, controlling for age, reading achievement, and pretest performance on the same measure. Error bars represent 1 standard error. $p < .05$; $^* p < .01$; $^{***} p < .001$.

$N = 100$) = 9.55, $p = .023$. To explore this unexpected result, we examined the significance of parameter estimates versus the reference group (varying the reference group to examine each condition); these post hoc comparisons were not part of our preregistered plan. They showed that the hybrid number line training led to higher accuracy on consistent items (adjusted $M = .95$, $SE = .02$) than the pure number line training (adjusted $M = .86$, $SE = .03$; $p = .003$, $d = 0.88$), the square number line training (adjusted $M = .88$, $SE = .03$; $p = .009$, $d = 0.79$), and the square area model training (adjusted $M = .90$, $SE = .02$; $p = .031$, $d = 0.61$).

For ambiguous magnitude comparison items, the binomial regression revealed a significant effect of condition, $Wald(3, N = 100) = 40.03$, $p < .001$. We tested condition differences by examining the significance of parameter estimates versus the reference group (varying the reference group to examine each condition). On the ambiguous magnitude comparison items, participants in the hybrid unidimensional number line condition (adjusted $M = .61$, $SE = .03$) significantly outperformed those in the square number line condition (adjusted $M = .42$, $SE = .04$; $p < .001$, $d = 1.19$) and in the area model condition (adjusted $M = .34$, $SE = .03$; $p < .001$, $d = 1.75$) but did not significantly differ from those in the pure number line condition (adjusted $M = .53$, $SE = .03$; $p = .076$, $d = 0.51$). Performance in the area model condition was significantly worse than in the pure unidimensional number line condition ($p < .001$, $d = 1.20$) but did not significantly differ from performance in the square number line condition ($p = .089$, $d = 0.50$).

Exploratory analyses

We conducted exploratory analyses to address two unanticipated sources of potential bias in our preregistered analyses. First, our control measure of reading achievement had a higher than anticipated amount of missing data, leading to loss of data due to listwise deletion. Therefore, we excluded reading achievement as a control measure; we still included age and pretest performance as covariates. Rerunning our main analyses without reading achievement as a covariate resulted in the same pattern of results as shown in Fig. 2.

Second, our preregistration did not account for experimenter error during training. After data collection, we coded the training packets to determine whether the written product of the experimenter's demonstrations of the fractions $1/2$, $1/4$, and $1/6$ adhered to the training protocol. We coded errors in any of the three steps of the demonstration procedure (segment, shade, and place); coders did not code sessions where they had been the experimenters. Failure to shade (3.2% of experimenter demonstration trials) and failure to place (6.8% of trials) were the most common experimenter errors. Failure to segment (0.2% of trials), failure to do any steps (0.2% of trials), and other errors (2.9% of trials) also occurred. These errors were not evenly distributed across conditions; notably, failure to shade occurred only in the pure unidimensional number line condition and affected 6 participants. To estimate the effects of the training when conducted with high fidelity to the training protocol, a second set of exploratory analyses included only participants with no experimenter demonstration errors ($n = 117$). These analyses again excluded the reading achievement measure.

We examined our main outcome measure, transfer to fraction magnitude comparison, within this subsample for whom training fidelity was high. An ANCOVA on posttest magnitude comparison accuracy, controlling for age and pretest magnitude comparison accuracy, revealed a significant effect of condition, $F(3, 102) = 4.24$, $p = .007$, $\eta_p^2 = .111$ (Fig. 3). We conducted pairwise contrasts on the estimated marginal means of posttest magnitude comparison accuracy in each condition. Consistent with our main analysis, the hybrid unidimensional number line condition (adjusted $M = .56$, $SE = .03$) significantly outperformed the square number line condition (adjusted $M = .47$, $SE = .03$; $p = .020$, $d = 0.62$) and area model condition (adjusted $M = .43$, $SE = .03$; $p = .002$, $d = 0.89$) and did not significantly differ from the pure unidimensional number line (adjusted $M = .54$, $SE = .03$; $p = .552$, $d = 0.17$). In addition, and different from our main analysis, the pure unidimensional number line condition outperformed the area model condition ($p = .018$, $d = 0.72$).

Analyses of number line estimation and area model estimation within this subsample, and without reading achievement as a covariate, yielded results parallel to those in the preregistered analyses. These exploratory results suggest that the fraction magnitude comparison performance of the pure unidimensional number line condition in the preregistered analyses may have been lowered due to

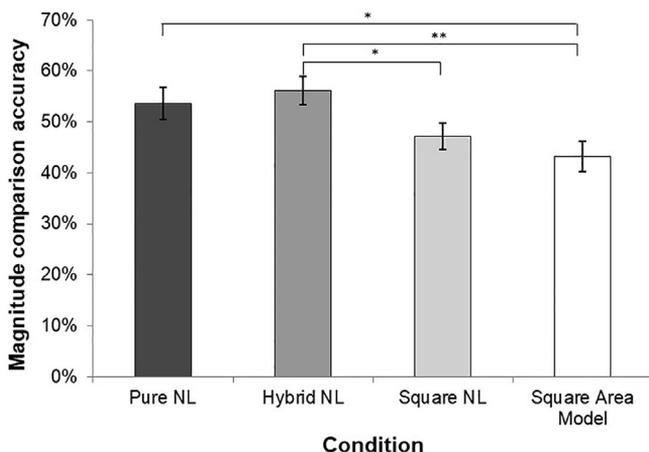


Fig. 3. Posttest fraction magnitude comparison accuracy by condition in exploratory analysis among participants with no experimenter errors in demonstration. Values are adjusted means controlling for age and pretest accuracy. Error bars represent 1 standard error. NL, number line. * $p < .05$; ** $p < .01$.

the presence of experimenter errors (i.e., failure to shade) in this condition. For the most part, however, these exploratory analyses confirmed the robustness of the results in the main, preregistered analyses.

Discussion

Results from this experimental training study provide causal evidence of a connection between fraction number line estimation and fraction magnitude concepts. Teaching young children to represent fractions on a number line led to higher performance on a transfer task—fraction magnitude comparisons—than a matched training using area models, replicating Hamdan and Gunderson (2017b). Two new experimental training conditions allowed us to tease apart specific features of the number line that led to this outcome. A critical new training condition—square number line training—involved training with a number line that was not unidimensional but did have left-to-right orientation and labeled endpoints (see Fig. 1C). Children performed worse after training with the square number line than after training with either of the unidimensional number lines (Fig. 1A and B). This is consistent with the theory that unidimensionality is a necessary feature of the number line for promoting transfer to fraction magnitude concepts.

Why is unidimensionality so important? We argue that the number line's unidimensionality is beneficial for learning because it aligns with a conceptual feature of real number magnitudes—that they can be ordered on a single dimension (Siegler & Lortie-Forgues, 2014). A unidimensional external number line representation is also well aligned with a mature mental representation of numerical magnitudes (Dehaene, 1992). We note that, under our Hypothesis 1 (a strict hypothesis favoring unidimensionality), we would have expected the pure unidimensional number line to outperform the hybrid unidimensional number line, which would in turn outperform the square number line and square area model conditions. This hypothesis was partially supported. The hybrid number line consistently outperformed both square conditions. However, the two unidimensional conditions performed similarly in terms of transfer to magnitude comparison. Whereas the hybrid unidimensional number line seemed to show (nonsignificantly) higher outcomes in our preregistered analyses (Fig. 2C), the two conditions looked nearly identical after accounting for experimenter error (Fig. 3). This suggests that the conceptual benefits of unidimensionality were preserved in both the pure and hybrid number line conditions. The lack of a difference between the pure and hybrid number line conditions was also inconsistent with our Hypothesis 3 (that the pure unidimensional number line would decrease performance by focusing children's attention on hash marks instead of spaces).

Furthermore, the square number line training did not lead to higher performance on fraction magnitude comparisons than the area model training (Fig. 1D), suggesting that the number line's features of left-to-right orientation and labeled endpoints—in the absence of unidimensionality—were not sufficient to promote transfer (in contrast to alternative Hypothesis 2). This is striking given the strength and early development of the association of left-to-right space with smaller-to-larger magnitudes (McCrink & Opfer, 2014; Rugani & de Hevia, 2017). However, it is important to note that left-to-right orientation and labeled endpoints may still be *necessary* for the number line to be beneficial. We chose not to test number line training conditions where either left-to-right orientation or labeled endpoints were absent because strong evidence already supports the importance of the number line's left-to-right orientation (e.g., Ebersbach, 2015). Indeed, it is possible that the two-dimensionality of the square number line led to lower performance in part by detracting from children's ability to focus on the number line's left-to-right orientation. Interestingly, the two-dimensionality of the square number line training led to lower performance despite the fact that the vertical dimension was irrelevant to the task. This pattern of results raises questions about how “thick” a hybrid number line can be and still maintain its benefits. Future research could examine parametric variations in the hybrid number line's height-to-width ratio to determine at what ratio children no longer benefit from this representation.

In addition to transferring to children's fraction magnitude comparison performance, the training conditions affected our other dependent variables, number line and area model estimation. First, number line estimation performance at posttest was greater in both unidimensional number line conditions (pure and hybrid) than in the square number line and area model training conditions. The lack of improvement on number line estimation in the square number line condition is particularly striking and suggests that the two-dimensionality of the training representation detracted from children's ability to view it as a number line at all.

On area model estimation, children in the pure unidimensional number line training condition performed worse than children in the other three conditions, which did not significantly differ from one another. This suggests that the slight two-dimensionality of the hybrid number line may have been especially beneficial, giving students an effective method for estimating area models while still maintaining the benefits of unidimensionality.

In addition to Hypotheses 1–3, which focused on the two new training conditions, we also preregistered three additional hypotheses. Hypothesis 4 stated that we expected to conceptually replicate the main findings of Hamdan and Gunderson (2017b). This hypothesis was mainly supported; as predicted, when compared with children trained using the area model, children trained in the hybrid unidimensional number line performed better at posttest in fraction number line estimation and magnitude comparison. However, our results did not support the second part of Hypothesis 4; we did not find that children trained in the area model outperformed children trained in the hybrid number line on area model estimation. Rather, both groups improved in area model estimation and did not significantly differ from one another at posttest. This finding is consistent with prior work and suggests that training using the hybrid number line may transfer not only to magnitude comparison but also to area model estimation (Hamdan & Gunderson, 2017b).

Hypothesis 5 predicted that the impact of condition on fraction magnitude comparisons would be significant for ambiguous and inconsistent items but not for consistent items. This was partially supported; the condition difference was significant for ambiguous items, and we found an unexpected significant impact of condition on consistent items. In contrast, Hamdan and Gunderson (2017b) found a significant training effect on ambiguous and inconsistent items but not on consistent items. Thus, across studies, these results are most robust for ambiguous items. Notably, in both the current study and prior work (Hamdan & Gunderson, 2017b), students performed well below chance on inconsistent items even at posttest. Inconsistent items may require substantial inhibitory control resources to avoid a prepotent, whole-number-based response, and a single 15-min training session was insufficient to overcome this whole-number bias. In contrast, ambiguous items reduce the need to inhibit whole-number-based strategies, which may allow children to more easily deploy new strategies taught during training.

Finally, Hypothesis 6 predicted that number line training would be beneficial for estimating unit fractions on the number line but not for estimating non-unit fractions and that there would be no dif-

ferential effect of condition on unit versus non-unit fraction area model estimation. The results did not support this prediction, and we found no significant differences for unit versus non-unit fractions. This suggests that the unidimensional number line trainings had a relatively broad impact on children's fraction concepts.

One limitation of this study is that, in the square number line training condition, what the experimenter called a "number line" (the square number line) may have conflicted with children's preexisting notions of what a number line should look like, potentially lowering performance. We called the square number line a "number line" to match the other number line conditions and to increase the likelihood that children would benefit from the properties it shares with the other number line conditions (i.e., labeled endpoints, left-to-right orientation). Children in our study were capable of learning from the square number line training, as shown by their significant improvement from pretest to posttest in area model estimation. Nevertheless, the lack of improvement in this condition on number line estimation and magnitude comparison may have stemmed, in part, from confusion about applying the label "number line" to a square number line. Future work could test whether the square number line representation would be more effective if not referred to as a "number line."

Our study tested the impact of these representations only among second- and third-graders. Thus, it is possible that the effects we found are specific to this age, when children have very little prior knowledge of fractions and when much of that prior knowledge is based on the area model representation. The number line representation may be especially helpful for learning about fraction magnitudes at this early stage, whereas other representations (or connecting different representations) may be more impactful for children who have greater fraction experience.

In addition, it would be beneficial to replicate these results using a design where experimenters are unaware of children's training condition to ensure that experimenter bias or cuing effects could not affect the results. However, in this study fraction magnitude comparison was collected directly from children using a computerized task, reducing the possibility of experimenter bias affecting this measure.

Furthermore, our study involved only a brief intervention (~15 min), and we believe that the presence of significant effects after such a brief time speaks to the power of these representations. Nevertheless, we note that children's performance on the magnitude comparison task remained poor at posttest. In particular, performance on items inconsistent with children's whole-number bias was well below chance and did not improve with training in any condition. This suggests that children's default mode of comparing fractions, even after training, was to compare the whole-number magnitudes of each fraction's components (numerators or denominators). Unidimensional number line training may give children new strategies for reasoning about fractions' overall magnitudes (e.g., visualizing locations and lengths along a number line) but might not directly reduce children's whole-number bias. Only when this whole-number bias was neutralized (on ambiguous magnitude comparison items) did unidimensional number line training lead to large improvements in fraction magnitude comparisons. These improvements are compelling given the short duration of the training and children's inexperience with fractions in second and third grades. It remains an open question whether a similar intervention with more training sessions would also help children to overcome their whole-number bias. Future research could investigate this and test whether the effects are durable beyond the immediate posttest assessed here.

Another limitation of the current study is the unanticipated level of experimenter error on our reading control measure and experimenters' fidelity to the training itself. However, the results of our preregistered analyses and exploratory analyses (adjusting for these issues) were very similar. Performance on the transfer task was significantly better in the hybrid unidimensional number line condition than in the square number line and square area model conditions. The key difference between our preregistered and exploratory analyses was in the interpretation of the results of the pure unidimensional number line condition. Only in our exploratory analyses, adjusting for experimenter error in training, was transfer performance significantly better in the pure unidimensional number line condition than in the area model condition (Fig. 3). This may be because in the pure unidimensional number line condition the experimenter in some cases failed to shade from zero to the fraction's location. This raises the fascinating possibility that shading the number line was a key aspect of the training's effectiveness. Shading the number line provides congruent continuous magnitude cues—both tempo-

ral (amount of time taken to shade) and spatial (amount of space shaded)—which may add to children’s magnitude understanding beyond the location information provided by the hash mark itself. The ability to process ratios of non-symbolic magnitudes, such as lengths, is related to symbolic fraction knowledge (Matthews, Lewis, & Hubbard, 2016), and our training may have been successful in part because it made explicit connections between the non-symbolic ratio processing system (i.e., shaded length and total length of the number line) and symbolic fractions.

This study represents a major step forward, showing that a short intervention can “move the needle” and thereby demonstrating a causal relation between the external number line representation and children’s fraction magnitude concepts. Future research could include more intensive intervention procedures and longer term outcome measures to determine whether the benefits are retained over time. In addition, future studies could test whether other features of the number line, such as continuity (vs. a discretized number line) and linearity (vs. a circular number line, similar to research on whole number lines; Siegler & Ramani, 2009), are necessary or sufficient for promoting transfer to fraction magnitude concepts. Another fruitful direction may be to investigate whether number lines are effective at promoting magnitude concepts for fractions greater than 1. With a number line, it is easy to extend the representation to include mixed numbers and improper fractions, whereas representing these with an area model is less intuitive and requires introducing new shapes. This may lead children to adopt the incorrect heuristic that fractions are always smaller than 1 (Kallai & Tzelgov, 2009). Thus, number line training may be particularly helpful for learning about improper fractions, which can be especially difficult for children to understand (Resnick et al., 2016).

In sum, this study bolsters the causal evidence that number line estimation promotes fraction magnitude concepts in young children. These data also provide insight into *why* the number line is beneficial for learning fraction magnitudes by specifying that unidimensionality is a critical feature. These results, in combination with other recent research (e.g., Fuchs et al., 2013; Hamdan & Gunderson, 2017b), have important implications for translational research using number lines in educational materials. They suggest that the assumption that area models are easier for young children to understand may be incorrect and that introducing fractions using number line representations in early elementary school may be an effective method for improving children’s fraction understanding.

Acknowledgments

This research was supported by a National Science Foundation CAREER Award (DRL-1452000) to Elizabeth Gunderson. We thank the children and teachers who gave their time to this research and the research assistants who helped to carry it out: Amma-Sika Adomako, Cory Ardekani, Tyler Burger, Alysa Cannon, Lillian Ham, Elizabeth Kohlbrenner, Sania Latif, Ying Lin, Kyle McCloskey, R.J. Nair, and Anza Thomas.

Appendix A. Training scripts

A.1. Number line training script

Keep in mind the following notes while administering the intervention:

1. x/y represents the fraction currently being instructed. Always say the fraction x/y as “one half”, “one third”, “two sixths”, etc., unless otherwise stated (i.e., where the script says “x over y,” you should instead say “one over two,” “one over three,” “two over six,” etc.).
2. Be attentive of child’s engagement in/attention to the task. If you notice child’s eyes diverting, or child becomes restless through his or her behavior or lack of attention, engagement, or delayed responses, suggest a break like this: “How about we take a break? Let’s stand up and get our wiggles out and then sit for a little bit!” Do this for up to 3 min, depending on how much time the child needs to regain enthusiasm and engagement in the task.
3. Throughout this task, participants will always be shown an example of the correct answer (remediation) whether they are correct in their demonstrations or not. Participants will be told “That’s

right” and shown an image of the right answer that matches their answer, or, they will be told “Actually” and shown the correct answer, given a new appropriate testing sheet (depending on where they are in the task), and asked to make their work look like the example shown. If children are not able to approximately replicate the correct answer after seeing it, children are to be given a correct example to continue their progression throughout the task.

4. Sections of this lesson that are labeled (segment, shade, place) correspond to a specific remediation protocol/figure with the same label. Present all children with remediation image and follow lesson according to whether they answer respective questions correctly or incorrectly.

A: Experimenter-led instruction
Introduction

“Today we are going to be playing a game where we are going to learn about fractions. Remember, you can tell me anytime if you do not want to play anymore after we start our game, okay? All right, let’s start!”

Show the unsegmented number line. Say: “This is a number line. This number line goes from 0 at this end to 1 at this end [say as you point to each endpoint on number line]. We can use this number line to show fractions.”

Say: “Fractions have a top part and a bottom part, like this number, x over y . This is a fraction. It has a number on top and a number on the bottom. We can call this fraction x over y because the x is on top and the y is on the bottom [say while pointing to appropriate part of the fraction]. We can also call this fraction x/y (e.g., one half). Let’s show this fraction using the number line we saw before.”

“First, we need to look at the bottom number [say while pointing to denominator]. The number on the bottom tells us how many equal parts we need to make on the number line.”

Step 1: Segment

Say: “So, we need to make y equal parts on the number line like this [do segmentation on number line and count parts afterward while pointing to each part]. See, one, two, etc.!”

Step 2: Shade

Say: “After we see how many equal parts we need on the number line, we need to look at the top number [say while pointing to the numerator]. This number tells us how many equal parts we need to color, starting from zero. So, we need to color x equal part(s) like this” [highlight over appropriate part of number line, and afterward count part(s) highlighted while pointing to parts highlighted].

Step 3: Place

Say: “If someone asks where x/y is on the number line, we should say here [point to the right endpoint of shaded part]. We can show it by drawing a hash mark at the end of the part we colored like this [draw hash mark]. We should write x/y above the hash mark like this [write fraction above hash mark location]. Remember, we can call this x over y or x/y . See how we showed x/y on the number line!?”

B. Student practice, segmented line

(Show segmented number line and fraction.) “Now, look at this number line. The number line is already divided into y equal parts.”

Step 1: Shade

Say: “We want to show x/y . How many equal parts do we need to color starting from zero?”

We need to look at the number on the top to know how many equal parts we need to color starting from zero [say while pointing to numerator]. Let's show x/y on this number line. Can you please color the right number of parts?"

—Say if child answers correctly: "That's right! To show x/y , you colored x equal parts on the number line and it looks like this" [show child appropriate color remediation page]."

—Say if child answers incorrectly: "Actually, to show x/y , you need to color x equal parts on the number line because the x is the number on top [say while pointing to numerator] and it tells us how many equal parts to color, starting from zero. See, like this [show child appropriate color remediation page]. Can you make this number line [present child with appropriate new worksheet for the fraction color task] look like this one? [say while pointing to color remediation page]. Great!" [continue in this way even if child is still unable to replicate correct answer].

Step 2: Place

Say: "Great! Now, can you put x/y where it goes on this number line?"

—Say if child answers correctly: "That's right! You show x/y by drawing a hash mark at the end of the part you colored and writing the fraction above the hash mark. It looks like this [show child appropriate place remediation page]. Great!"

—Say if child answers incorrectly: "Actually, you show x/y by drawing a hash mark at the end of the part you colored and writing the fraction above the hash mark. See, like this [show child appropriate place remediation page]. Now, can you make this number line [present child with appropriate new worksheet for the fraction place task] look like this one? [show place remediation page from before]. Great!" [continue in this way even if child is still unable to replicate correct answer].

C. Student practice, unsegmented line

[Show unsegmented number line]

Step 1: Segment

Say: "Now, look at this number line. Let's show x/y on this number line. First, we need to know how many equal parts we should make on this number line. How many equal parts do we need to make? We need to look at the number on the bottom [say while pointing to denominator] to know how many equal parts we need to make on the number line. Can you please make the right number of equal parts?"

—Say if child answers correctly: "That's right! To show x/y , you start by making y equal parts on the number line, see, one, two, etc. [say while pointing to the segments made by the child] and it looks like this [show appropriate segment remediation page]. Great!"

—Say if child answers incorrectly: "Actually, to show x/y , you start by making y equal parts on the number line. Remember, you should look at the number on the bottom [say while pointing to denominator] to tell you how many equal parts you need to make. See, like this [show appropriate segment remediation page]. See how this number line has y equal parts? Can you make this number line [present child with appropriate new worksheet for that fraction segment task] look like this one? [say while pointing to the segment remediation page]. Great!" [continue in this way even if child is still unable to replicate correct answer].

Step 2: Shade

Say: "We want to show x/y . How many equal parts do we need to color starting from zero?"

We need to look at the number on the top to know how many equal parts we need to color starting from zero [say while pointing to numerator]. Let's show x/y on this number line. Can you please color the right number of parts?"

—Say if child answers correctly: "That's right! To show x/y , you colored x equal parts on the number line and it looks like this [show child appropriate color remediation page].

—Say if child answers incorrectly: “Actually, to show x/y , you need to color x equal parts on the number line because the x is the number on top [say while pointing to numerator] and it tells us how many equal parts to color, starting from zero. See, like this [show child appropriate color remediation page]. Can you make this number line [present child with appropriate new worksheet for that fraction color task] look like this one? [say while pointing to color remediation page]. Great!” [continue in this way even if child is still unable to replicate correct answer].

Step 3: Place

Say: “Great! Now, can you put x/y where it goes on this number line?”

—Say if child answers correctly: “That’s right! You show x/y by drawing a hash mark at the end of the part you colored and writing the fraction above the hash mark. It looks like this [show child appropriate place remediation page]. Great!”

—Say if child answers incorrectly: “Actually, you show x/y by drawing a hash mark at the end of the part you colored and writing the fraction above the hash mark. See, like this [show appropriate place remediation page]. Now, can you make this number line [present child with appropriate new worksheet for that fraction place task] look like this one? [show place remediation page from before]. Great!” [continue in this way even if child is still unable to replicate correct answer].

Continue through this script again for the remaining training fraction families ($x/4$, $x/6$).

At end of training, say: “Now we can use what we’ve learned here to help us answer some questions.”

A.2. Area model training script

Keep in mind the following notes while administering the intervention:

1. x/y represents the fraction currently being instructed. Always say the fraction x/y as “one half,” “one third,” “two sixths,” etc., unless otherwise stated (i.e., where the script says “ x over y ,” you should instead say “one over two,” “one over three,” “two over six,” etc.).

2. Be attentive of child’s engagement in/attention to the task. If you notice child’s eyes diverting, or the child become restless through his or her behavior or lack of attention, engagement, or delayed responses, suggest a break like this: “How about we take a break? Let’s stand up and get our wiggles out and then sit for a little bit!” Do this for up to 3 min, depending on how much time the child needs to regain enthusiasm and engagement in the task.

3. Throughout this task, participants will always be shown an example of the correct answer (remediation) whether they are correct in their demonstrations or not. Participants will be told “That’s right” and shown an image of the right answer that matches their answer, or they will be told “Actually” and shown the correct answer, given a new appropriate testing sheet (depending on where they are in the task), and asked to make their work look like the example shown. If children are not able to approximately replicate the correct answer after seeing it, children are to be given a correct example to continue their progression throughout the task.

4. Sections of this lesson that are labeled (segment, shade, place) correspond to a specific remediation protocol/figure with the same label. Present all children with remediation image, and follow lesson according to whether they answer respective questions correctly or incorrectly.

Always start coloring square from the top right and proceed clockwise.

A: Experimenter-led instruction

Introduction

“Today we are going to be playing a game where we are going to learn about fractions. Remember, you can tell me anytime if you do not want to play anymore after we start our game, okay? All right, let’s start!”

[Show unsegmented square]. Say: “This is one square [say as you trace square with fingers]. We can use this square to show fractions.”

Say: "Fractions have a top part and a bottom part like this number, x over y . This is a fraction. It has a number on top and a number on the bottom. We can call this fraction x over y because the x is on top and the y is on the bottom [say while pointing to appropriate part of the fraction]. We can also call this fraction x/y (e.g., one half). Let's show this fraction using the square we saw before."

"First, we need to look at the bottom number [say while pointing to denominator]. The number on the bottom tells us how many equal parts we need to make on the number line."

Step 1: Segment

Say: "So, we need to make y equal parts on the square like this [do segmentation on square and count parts afterward while pointing to each part]. See, one, two, etc.!"

Step 2: Shade

Say: "After we see how many equal parts we need on the square, we need to look at the top number [say while pointing to the numerator]. This number tells us how many equal parts we need to color. So, we need to color x equal part(s) like this" [color appropriate part of square, and afterward count part(s) colored while pointing to parts colored].

Step 3: Place

Say: "If someone asks where x/y is on the square, we should say here [point to the middle of the shaded part]. We can show it by writing the fraction next to the part we colored like this [write fraction outside of colored part to the right of shape]. We should draw an arrow like this [draw arrow to the left of fraction]. Remember, we can call this x over y or x/y . See how we showed x/y on the square?"

B. Student practice, segmented square

[Show segmented square and fraction]. Say: "Now, look at this square. The square is already divided into y equal parts.

Step 1: Shade

Say: "We want to show x/y . How many equal parts do we need to color?"

"We need to look at the number on the top to know how many equal parts we need to color [say while pointing to numerator]. Let's show x/y of this square. Can you please color the right number of parts?"

—Say if child answers correctly: "That's right! To show x/y , you colored x equal parts on the square and it looks like this" [show child appropriate color remediation page].

—Say if child answers incorrectly: "Actually, to show x/y , you need to color x equal parts on the square because the x is the number on top [say while pointing to numerator] and it tells us how many equal parts to color. See, like this [show child appropriate color remediation page]. Can you make this square [present child with appropriate new worksheet for that fraction color task] look like this one? [say while pointing to color remediation page]. Great!" [continue in this way even if child is still unable to replicate correct answer].

Step 2: Place

Say: "Great! Now, can you put x/y where it goes on this square?"

—Say if child answers correctly: "That's right! You show x/y by writing the fraction next to the part you colored and drawing an arrow pointing to the part you colored. It looks like this [show child appropriate place remediation page]. Great!"

—Say if child answers incorrectly: "Actually, you show x/y by writing the fraction next to the part you colored and drawing an arrow pointing to the part you colored. See, like this [show child appro-

ropriate place remediation page]. Now, can you make this square [present child with new worksheet for that fraction place task] look like this one? [show place remediation page from before]. Great!" [continue in this way even if child is still unable to replicate correct answer].

C. Student practice, unsegmented square

[Show unsegmented square]. Say: "Now, look at this square. Let's show x/y of this square.

Step 1: Segment

"First, we need to know how many equal parts we should make on this square. How many equal parts do we need to make?"

"We need to look at the number on the bottom [say while pointing to denominator] to know how many equal parts we need to make on the square. Can you please make the right number of equal parts?"

—Say if child answers correctly: "That's right! To show x/y , you start by making y equal parts on the square, see, one, two, etc. [say while pointing to the segments made by the child] and it looks like this [show appropriate segment remediation page]. Great!"

—Say if child answers incorrectly: "Actually, to show x/y , you start by making y equal parts on the square. Remember, you should look at the number on the bottom [say while pointing to denominator] to tell you how many equal parts you need to make. See, like this [show appropriate segment remediation page]. See how this square has y equal parts? Can you make this square [present child with appropriate new worksheet for that fraction segment task] look like this one? [say while pointing to segment remediation page]. Great! [continue in this way even if child is still unable to replicate correct answer].

Step 2: Shade

Say: "We want to show x/y . How many equal parts do we need to color? We need to look at the number on the top to know how many equal parts we need to color [say while pointing to numerator]. Let's show x/y of this square. Can you please color the right number of parts?"

—Say if child answers correctly: "That's right! To show x/y , you colored x equal parts on the square and it looks like this" [show child appropriate color remediation page].

—Say if child answers incorrectly: "Actually, to show x/y , you need to color x equal parts on the square because the x is the number on top [say while pointing to numerator] and it tells us how many equal parts to color. See, like this [show child appropriate color remediation page]. Can you make this square [present child with appropriate new worksheet for that fraction color task] look like this one? [say while pointing to color remediation page]. Great!" [continue in this way even if child is still unable to replicate correct answer].

Step 3: Place

Say: "Great! Now, can you put x/y where it goes on this square?"

—Say if child answers correctly: "That's right! You show x/y by writing the fraction next to the part you colored and drawing an arrow pointing to the part you colored. It looks like this [show child appropriate place remediation page]. Great!"

—Say if child answers incorrectly: "Actually, you show x/y by writing the fraction next to the part you colored and drawing an arrow pointing to the part you colored. See, like this [show child appropriate place remediation page]. Now, can you make this square [present child with new worksheet for that fraction place task] look like this one? [show place remediation page from before]. Great!" [continue in this way even if child is still unable to replicate correct answer].

Continue through this script again for the remaining training fraction families ($x/4$, $x/6$). At end of training, say: "Now we can use what we've learned here to help us answer some questions."

Appendix B. Fraction magnitude comparison items

See [Table B1](#).

Table B1
Magnitude comparison items.

Item	Bigger whole number strategy category	Pretest accuracy	Posttest accuracy	In larger fraction, numerator is ...	In larger fraction, denominator is ...
1/4 vs. 4/5	Consistent	.89	.91	Larger	Larger
2/5 vs. 4/5	Consistent	.85	.86	Larger	Equal
2/5 vs. 6/7	Consistent	.89	.88	Larger	Larger
3/4 vs. 1/2	Consistent	.92	.92	Larger	Larger
3/4 vs. 5/6	Consistent	.92	.90	Larger	Larger
4/7 vs. 2/6	Consistent	.84	.83	Larger	Larger
6/7 vs. 3/5	Consistent	.85	.85	Larger	Larger
6/7 vs. 5/7	Consistent	.83	.82	Larger	Equal
1/2 vs. 2/5	Inconsistent	.10	.12	Smaller	Smaller
1/3 vs. 1/6	Inconsistent	.07	.13	Equal	Smaller
1/4 vs. 1/2	Inconsistent	.06	.08	Equal	Smaller
1/5 vs. 1/2	Inconsistent	.10	.11	Equal	Smaller
2/3 vs. 3/6	Inconsistent	.08	.14	Smaller	Smaller
2/3 vs. 3/7	Inconsistent	.08	.14	Smaller	Smaller
2/5 vs. 2/4	Inconsistent	.09	.18	Equal	Smaller
3/5 vs. 3/4	Inconsistent	.07	.18	Equal	Smaller
1/7 vs. 2/3	Ambiguous	.25	.40	Larger	Smaller
1/7 vs. 2/6	Ambiguous	.51	.60	Larger	Smaller
2/3 vs. 1/4	Ambiguous	.49	.68	Larger	Smaller
2/3 vs. 1/5	Ambiguous	.34	.44	Larger	Smaller
2/3 vs. 1/6	Ambiguous	.26	.42	Larger	Smaller
2/4 vs. 1/5	Ambiguous	.47	.55	Larger	Smaller
2/4 vs. 1/6	Ambiguous	.33	.49	Larger	Smaller
2/6 vs. 3/4	Ambiguous	.32	.38	Larger	Smaller
2/6 vs. 3/5	Ambiguous	.44	.58	Larger	Smaller
2/7 vs. 3/5	Ambiguous	.37	.45	Larger	Smaller
3/7 vs. 4/6	Ambiguous	.48	.49	Larger	Smaller
4/5 vs. 3/7	Ambiguous	.34	.45	Larger	Smaller

Appendix C. Supplemental analyses

Fraction addition task: To explore whether training would transfer to fraction addition skill, participants completed a four-item fraction addition task at both pretest and posttest. The four items all included adding fractions with like denominators: $4/7 + 2/7$, $2/6 + 3/6$, $2/4 + 1/4$, and $1/3 + 1/3$. The items were presented on paper, and children were asked to write their responses. [Table C2](#) reports mean accuracy on this task, which was quite low, and tests of raw condition differences and pretest to posttest change.

We also sought to examine the effect of condition on fraction addition. Parallel to our main preregistered analyses, we conducted a binomial regression on posttest fraction addition accuracy, controlling for age and reading achievement. It was not possible to control for pretest fraction addition accuracy because of floor performance at pretest in one condition. This analysis showed no main effect of condition on posttest fraction addition accuracy, $Wald(3, N = 104) = 6.36, p = .095$.

In a second binomial regression, parallel to the exploratory analyses in the main text, we removed reading achievement as a control and restricted the analysis to participants for whom there was no error in the experimenter demonstration. In this analysis, there was a significant main effect of condition on posttest fraction addition accuracy, $Wald(3, N = 114) = 8.02, p = .046$. With the area model as the reference group, the parameter estimates indicated that each of the three number line conditions

Table C1

Means (and standard deviations) of pretest to posttest change by condition.

Measure	Pure unidimensional NL training	Hybrid unidimensional NL training	Square NL training	Area model training
Number line estimation (PAE)	-.092 (.102) ^{***}	-.113 (.110) ^{***}	-.013 (.081)	-.001 (.083)
NL unit PAE	-.076 (.148) ^{**}	-.151 (.128) ^{***}	-.000 (.123)	-.031 (.160)
NL non-unit PAE	-.103 (.131) ^{***}	-.081 (.132) ^{**}	-.017 (.122)	-.014 (.078)
Area model estimation (PAE)	-.037 (.092) [*]	-.078 (.144) ^{**}	-.071 (.080) ^{***}	-.084 (.116) ^{***}
AM unit PAE	-.054 (.153)	-.105 (.155) ^{***}	-.047 (.104) [*]	-.115 (.175) ^{**}
AM non-unit PAE	-.023 (.106)	-.059 (.181) [*]	-.089 (.115) ^{***}	-.068 (.125) ^{**}
Magnitude comparison accuracy	.070 (.137) ^{**}	.129 (.221) ^{**}	.029 (.111)	.016 (.127)
MC consistent items accuracy	-.039 (.236)	.034 (.239)	-.011 (.133)	.024 (.257)
MC inconsistent items accuracy	.059 (.282)	.105 (.331)	.018 (.118)	.000 (.187)
MC ambiguous items accuracy	.151 (.291) ^{**}	.209 (.359) ^{**}	.062 (.229)	.021 (.243)

Note. Change score is posttest score minus pretest score. Negative scores on PAE indicate improvement because lower scores represent lower error. Tests of significance are paired-samples *t* tests comparing pretest with posttest performance. NL, number line; AM, area model; MC, magnitude comparison; PAE, percentage absolute error.

- ^{*} *p* < .05.
^{**} *p* < .01.
^{***} *p* < .001.

Table C2

Means (and standard deviations) of accuracy on fraction addition task by condition.

	Pure unidimensional NL training	Hybrid unidimensional NL training	Square NL training	Area model training	Test of condition difference
Pretest fraction addition accuracy	.00 (.00)	.03 (.16)	.10 (.29)	.02 (.13)	$F(3, 138) = 2.07$, $p = .107$, $\eta_p^2 = .043$
Posttest fraction addition accuracy	.09 (.28)	.16 (.37)	.14 (.36)	.03 (.17)	$F(3, 143) = 1.18$, $p = .318$, $\eta_p^2 = .023$
Pretest to posttest change in fraction addition accuracy	.09 (.28)	.14 (.34) [*]	.04 (.19)	.01 (.04)	NA

Note. Change score is posttest score minus pretest score. Tests of significance are paired-samples *t* tests comparing pretest with posttest performance. NL, number line; NA, not applicable.

- ^{*} *p* < .05.

had significantly higher posttest fraction addition accuracy than the area model ($ps < .05$). Rerunning the model with each of the number line conditions as the reference group indicated that none of the number line conditions significantly differed from each other. Although these results suggest that number line training may have benefits that transfer to fraction addition, we interpret these results with caution given that they were shown only in our exploratory analyses.

Appendix D. Estimation tasks: Coding guidelines and reliability

D.1. Number line coding

At both pretest and posttest, students labeled fractions on a 0-to-1 number line. Number line coding involved measuring the location on a 0-to-1 number line at which participants marked a fraction. Common instances of difficult to code items are mentioned below along with how we decided to code them. Blue lines and circles are used to indicate where we would code “marked” on the number line.

Child segments the line multiple times: When children segmented the line multiple times but did not indicate which mark they meant, the trial was coded as unscorable (Fig. D1A).

Hard to identify labels: When participants made multiple hash marks and then labeled one, we coded the one that they labeled. If it was unclear which hash mark was labeled, we coded the hash mark most directly under the written fraction or the arrow (Fig. D1B).

Mislabeled line: When students labeled the number line but used the wrong number, we ignored that the label was incorrect and coded the chosen line as the one below the label (Fig. D1C).

Scanned number lines were lined up with a ruler in Adobe Photoshop to determine the start point, the endpoint, and the mark point of each line. The mark point for each line was measured in millimeters and coded twice by separate coders. If the difference between the two measurements was greater than 1 mm, then the trial was flagged as a disagreement. In this study, 1 mm accounts for 0.57% of the total length of the number line used. Coders agreed on 90.64% of the 4144 trials. The remaining 9.36% of trials were checked by an independent experienced coder, who determined whether the first coder or second coder was correct.

D.2. Area model coding

In the area model portion of the pretest and posttest, we used Adobe Photoshop to determine the number of pixels shaded by the student for each trial. In doing so, we encountered a variety of shading styles that could be interpreted in varying ways. To remain reliable between coders, we reviewed ambiguous shading situations and came to conclusions as a unit. Below are the situations we encountered and our consensus on how to solve them. We have included sample shading situations. Blue outlines indicate the areas that we would measure as shaded.

Using Xs along with shading: When students shaded some segments of the square and crossed out other segments, we decided that this type of “crossing out” signified that these segments should not

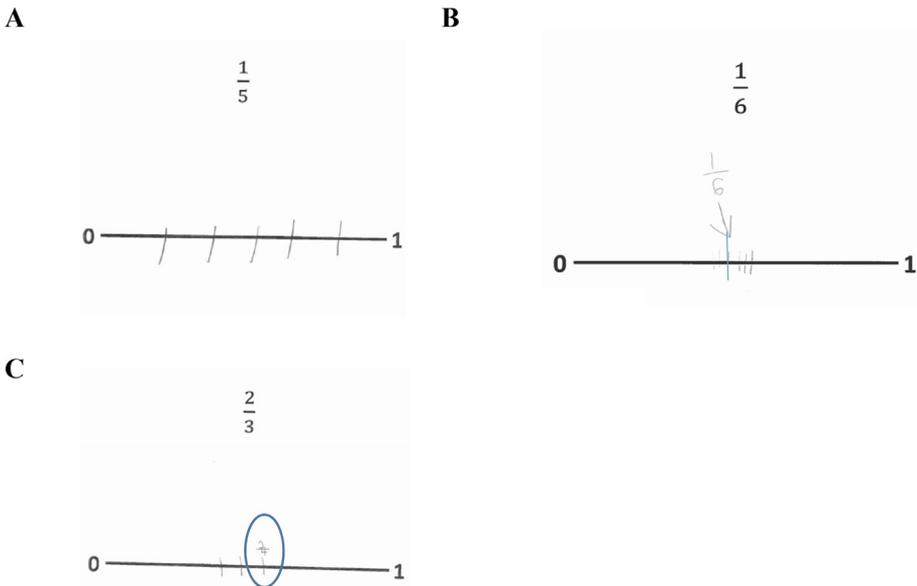


Fig. D1. Guidelines used when coding shaded number lines. These situations included students segmenting a line multiple times (A), providing hard to identify labels or arrows (B), and mislabeling the number line (C). Blue shapes indicate the hash marks that were coded as the demarcation. (For interpretation of the references to color in this figure legend and the text, the reader is referred to the Web version of this article.)

be counted as shading. In these cases, we found the pixel count of the shaded segments without including the “X’ed out” segments (Fig. D2A).

Items where children did not shade: We flagged and excluded data where students drew a picture or objects in an area model (Fig. D2B).

Items where students segmented the square but shaded only a portion of the segmented area: In cases where a portion of a segment was shaded, we made the assumption that students meant for the whole segment to be shaded. Therefore, we used the pixel count of the entire segmented area rather than just the shaded portion (Fig. D2C).

Items where students shaded only a portion of the square but did not segment: In these cases, because there was no segmentation, we measured the pixel count of the shaded area of the area model (Fig. D2D).

Minimal shading of a segment: We counted a segment as shaded even if the mark was minimal unless it was clearly an unintentional mark, for example, accidentally coloring over a segmentation line (Fig. D2E).

All area model trials were coded twice by separate coders. Measurements were done using Adobe Photoshop, and data were entered in Microsoft Excel. First and second entries of pixel counts were

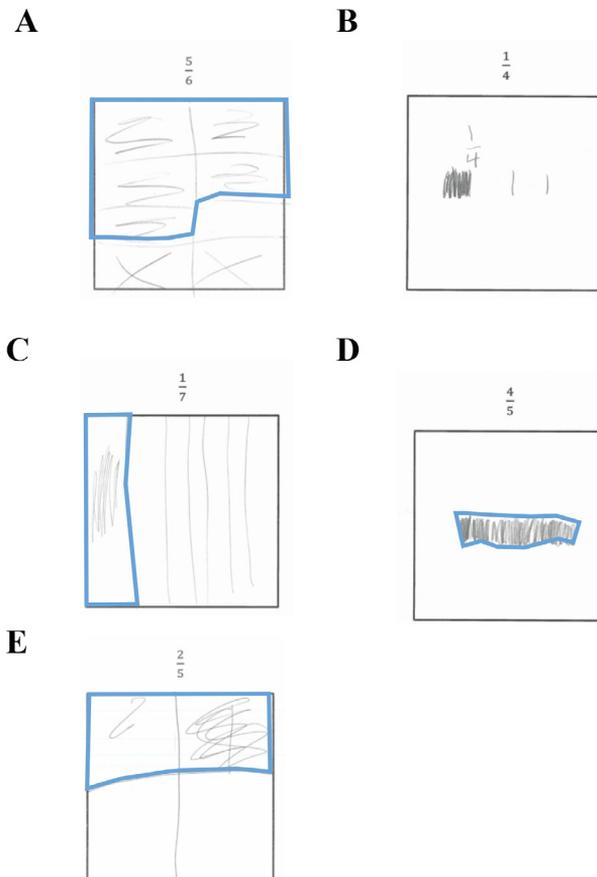


Fig. D2. Guidelines used when coding ambiguously shaded area models. These situations included shading some square segments and crossing others out (A), drawing pictures or number lines in the area model (B), shading small portions of a segmented space (C), shading without segmenting (D), and shading very small portions of a segment (E). Blue outlines indicate portion that was coded as shaded. (For interpretation of the references to color in this figure legend and the text, the reader is referred to the Web version of this article.)

compared by finding the absolute difference between the two entered values for number of pixels shaded. If the difference between the two measurements was greater than 100,000 pixels, then the trial was flagged as a disagreement (100,000 pixels account for 2.1% of total area of 18.4×18.4 cm rectangle [pixel area is 4,721,929] and 2.3% of total area shaded for 17.5×17.5 cm rectangle [pixel area is 4,272,489]). Of the 4144 trials coded, the experimenters agreed on 90.85% of trials. The remaining 9.15% of trials were checked by an independent experienced coder, who determined whether the first coder or second coder was correct.

References

- 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.
- Barth, H., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science*, *14*, 125–135.
- Booth, J. L., & Newton, K. J. (2012). Fractions: Could they really be the gatekeeper's doorman? *Contemporary Educational Psychology*, *37*, 247–253.
- Booth, J. L., Newton, K. J., & Twiss-Garrity, L. K. (2014). The impact of fraction magnitude knowledge on algebra performance and learning. *Journal of Experimental Child Psychology*, *118*, 110–118.
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42.
- Ebersbach, M. (2015). Evidence for a spatial-numerical association in kindergartners using a number line task. *Journal of Cognition and Development*, *16*, 118–128.
- 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.
- Hamdan, N., & Gunderson, E. A. (2017a). Number line as a representation for fraction magnitudes. Retrieved from <<https://osf.io/asgz/>>.
- Hamdan, N., & Gunderson, E. A. (2017b). The number line is a critical spatial-numerical representation: Evidence from a fraction intervention. *Developmental Psychology*, *53*, 587–596.
- 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.
- Hurst, M., & Cordes, S. (2016). Rational-number comparison across notation: Fractions, decimals, and whole numbers. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 281–293.
- luculano, T., & Butterworth, B. (2011). Understanding the real value of fractions and decimals. *Quarterly Journal of Experimental Psychology*, *64*, 2088–2098.
- Kallai, A. Y., & Tzelgov, J. (2009). A generalized fraction: An entity smaller than one on the mental number line. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1845–1864.
- Kloosterman, P. (2010). Mathematics skills of 17-year-olds in the United States: 1978 to 2004. *Journal for Research in Mathematics Education*, *41*, 20–51.
- Matthews, P. G., Lewis, M. R., & Hubbard, E. M. (2016). Individual differences in nonsymbolic ratio processing predict symbolic math performance. *Psychological Science*, *27*, 191–202.
- McCrink, K., & Opfer, J. E. (2014). Development of spatial-numerical associations. *Current Directions in Psychological Science*, *23*, 439–445.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core state standards*. Washington, DC: Author.
- Opfer, J. E., Thompson, C. A., & Furlong, E. E. (2010). Early development of spatial-numerical associations: Evidence from spatial and quantitative performance of preschoolers. *Developmental Science*, *13*, 761–771.
- Resnick, I., Jordan, N. C., Hansen, N., Rajan, V., Rodrigues, J., Siegler, R. S., & Fuchs, L. S. (2016). Developmental growth trajectories in understanding of fraction magnitude from fourth through sixth grade. *Developmental Psychology*, *52*, 746–757.
- Rinne, L. F., Ye, A., & Jordan, N. C. (2017). Development of fraction comparison strategies: A latent transition analysis. *Developmental Psychology*, *53*, 713–730.
- Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science*, *24*, 1301–1308.
- Rugani, R., & de Hevia, M.-D. (2017). Number-space associations without language: Evidence from preverbal human infants and non-human animal species. *Psychonomic Bulletin & Review*, *24*, 352–369.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.
- Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV tests of achievement*. Rolling Meadows, IL: Riverside.
- Shaki, S., Fischer, M. H., & Göbel, S. M. (2012). Direction counts: A comparative study of spatially directional counting biases in cultures with different reading directions. *Journal of Experimental Child Psychology*, *112*, 275–281.
- 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.
- 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., & Lortie-Forgues, H. (2014). An integrative theory of numerical development. *Child Development Perspectives*, *8*, 144–150.
- Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology*, *101*, 545–560.

- Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, *62*, 273–296.
- Slusser, E. B., & Barth, H. (2017). Intuitive proportion judgment in number-line estimation: Converging evidence from multiple tasks. *Journal of Experimental Child Psychology*, *162*, 181–198.
- Slusser, E. B., Santiago, R. T., & Barth, H. C. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology: General*, *142*, 193–208.
- 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.
- 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.
- Woodcock, R. W. (1999). What can Rasch-based scores convey about a person's test performance? In S. E. Embretson & S. L. Hershberger (Eds.), *The new rules of measurement: What every psychologist and educator should know* (pp. 105–127). Mahwah, NJ: Lawrence Erlbaum.