Improving Fraction Understanding in Sixth Graders With Mathematics Difficulties: Effects of a Number Line Approach Combined With Cognitive Learning Strategies

Christina A. Barbieri, Jessica Rodrigues, Nancy Dyson, and Nancy C. Jordan
University of Delaware

Abstract:
The effectiveness of an experimental middle school fraction intervention was evaluated. The intervention was centered on the number line and incorporated key principles from the science of learning. Sixth graders (N = 51) who struggled with fraction concepts were randomly assigned at the student level to the experimental intervention (n = 28) or to a business-as-usual control who received their school's intervention (n = 23). The experimental intervention occurred over 6 weeks (27 lessons). Fraction number line estimation, magnitude comparisons, concepts, and arithmetic were assessed at pretest, posttest, and delayed posttest. The experimental group demonstrated significantly more learning than the control group from pretest to posttest, with meaningful effect sizes on measures of fraction concepts (g = 1.09), number line estimation as measured by percent absolute error ($g = -.85$), and magnitude comparisons (g = .82). These improvements held at delayed posttest 7 weeks later. Exploratory analyses showed a significant interaction between classroom attentive behavior and intervention group on fraction concepts at posttest, suggesting a buffering effect of the experimental intervention on the normally negative impact of low attentive behavior on learning. A number line–centered approach to teaching fractions that also incorporates research-based learning strategies helps struggling learners to make durable gains in their conceptual understanding of fractions.

Keywords: fraction, fraction magnitude, intervention, mathematics difficulties, number line

Corresponding author: Christina Areizaga Barbieri, barbieri@udel.edu
University of Delaware School of Education
113 Willard Hall Education Building, Newark, DE 19716, USA

Full citation: Barbieri, C. A., Rodrigues, J., Dyson, N., & Jordan, N. C. (2019, June 20). Improving Fraction Understanding in Sixth Graders With Mathematics Difficulties: Effects of a Number Line Approach Combined With Cognitive Learning Strategies. Journal of Educational Psychology. Advance online publication. http://dx.doi.org/10.1037/edu0000384

Funding: Funding for this research was provided by the Institute of Education Sciences, U.S. Department of Education, Grants R324A160127 and R305A100150. IRB Project Name: Developing a Fraction Sense Intervention; IRBnet ID: 931086.

Educational Impact and Implications Statement

A mathematics intervention that used a number line– centered approach and validated learning principles to teach fraction concepts helped struggling sixth graders improve their fraction understanding. After participating in the intervention, students performed better on assessments of

fraction concepts, number line estimation, and magnitude comparisons than a group of students who received their school's regular intervention, and these improvements held seven weeks later. Findings suggest that students who are struggling with fractions, even after receiving several years of formal fraction instruction in school, can still make large gains in their understanding, preparing them for more advanced mathematics and for success in STEM related fields.

Introduction

A strong foundation in fractions helps students succeed in mathematics (National Mathematics Advisory Panel, 2008). More specifically, fraction *magnitude* understanding supports algebra proficiency (Booth & Newton, 2012). Because Algebra is a requisite skill for enrollment in advanced level mathematics courses as well as many science courses, algebra proficiency leads to a higher rate of college admittance, and eventual pursuit of careers in STEM disciplines (Chen, 2009; Matthews & Farmer, 2008; Schneider, Swanson, & Riegle-Crumb, 1998). Moreover, fraction knowledge is needed for many non-STEM jobs (Handel, 2016) as well as for everyday activities, such as managing money, cooking, and doing home repairs.

Unfortunately, many students enter sixth grade with a tenuous grasp of fractions, even after several years of instruction on the topic (Resnick et al., 2016). Students who enter seventh grade without foundational knowledge of fractions face cascading mathematics difficulties (Mazzocco & Devlin, 2008). To address this problem, we developed and evaluated an experimental intervention for entering sixth graders who exhibit low knowledge of fraction concepts. Our experimental intervention is informed by current research on fraction learning; specifically, the importance of understanding fractions as magnitudes that can be represented on the number line. Numerical magnitude knowledge uniquely predicts a range of mathematical competencies (see Schneider, Thompson, & Rittle-Johnson, 2018 for a review). Moreover, deficits in under-standing of symbolic numerical magnitudes characterize students with mathematics difficulties and disabilities (e.g., Butterworth, 2005; Butterworth & Reigosa-Crespo, 2007). *Fraction* magnitude knowledge is especially predictive of mathematics proficiency, beyond whole number skills and general cognitive competencies (Resnick et al., 2016). The number line is an effective but often underused tool for developing fraction magnitude knowledge in struggling students (Dyson, Jordan, Rodrigues, Barbieri, & Rinne, 2018; Fuchs et al., 2014; Gersten, Schumacher, & Jordan, 2017; Saxe, Diakow, & Gearhart, 2013). In addition to centering instruction on the number line, our approach aims to bolster students' fraction skills through application of research-based learning principles from cognitive science (Brown, Roediger, & McDaniel, 2014; Rittle-Johnson & Jordan, 2016).

Transitioning From Whole Numbers to Fractions

When learning fractions, students gradually expand their understanding of numbers; they take into account differences, as well as similarities, between fractions and whole numbers. Whole numbers are represented linearly with each number being exactly one more than the previous number and only one number represents each magnitude. Fractions, on the other hand, can be represented in countless ways (e.g., 1/5, 2/10, and so on). There are infinite fractional parts between integers, and fractions can be less than, equal to, or more than one (Resnick et al., 2016). When determining the magnitude of a fraction, the size of the numerator or denominator cannot be considered in isolation as in separate whole numbers. Larger numbers in the fraction do not always signify larger magnitudes (e.g., 1/4 < 1/2). With fractions less than one, multiplication does not lead to a product greater than a factor and division does not lead to a quotient smaller than a dividend.

Students often incorrectly apply whole number logic to fractions (DeWolf & Vosniadou, 2015; Siegler, Thompson, & Schneider, 2011; Vamvakoussi & Vosniadou, 2010), and the problem seems to be especially pervasive in low-achieving students (Malone & Fuchs, 2017). In a study of fraction arithmetic errors, it was found that low-achieving students focused on the size of the fractional parts rather than the relations among the parts (Malone & Fuchs, 2017). Importantly, these students' errors reflected poor magnitude understanding rather than difficulties with part–

whole knowledge more generally.

Reasoning About Fractions on the Number Line

Numerical magnitude reasoning is reflected by students' ability to estimate magnitudes on the number line. Understanding that all real numbers are represented as magnitudes on a number line provides a unifying framework for number learning (e.g., Siegler et al., 2011). For example, a fraction of 1/19 is very close to 0 relative to 6/7 which is closer to 1, and 5/4 is greater than 1. Students with stronger whole number estimation skills in third grade are more likely to perform better on fraction concepts and procedures measures in fourth and sixth grades (Bailey et al., 2015; Fuchs et al., 2013).

Fraction magnitude knowledge predicts both broad and more specific mathematics outcomes, over and above general cognitive abilities and whole number skills. A longitudinal study found that growth in fraction number line estimation (FNLE) acuity between fourth and sixth grades predicts mathematics achievement at the end of sixth grade, even when controlling for a constellation of domain general and domain specific abilities (Resnick et al., 2016). A troubling finding was that a significant number of students showed little to no growth in fraction number line estimation accuracy between fourth and sixth grade, even though they had received three years of fraction instruction in school. Fraction arithmetic is typically introduced in fourth grade with addition and subtraction and remains the primary focus through sixth grade (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). At this point, formal fraction arithmetic instruction typically comes to a close with division of fractions and students move on to prealgebra instruction. FNLE acuity specifically predicts prealgebra skills (Booth & Newton, 2012), a key element of sixth-grade achievement (Booth, Newton, & Twiss-Garrity, 2014). One proposed explanation is that understanding rules that govern the relationship between the numerator and denominator can later be translated to reasoning about algebraic equations (e.g., “fractional representations of 1/3 will fit into the equation, $x/y = 1/3$, which is equivalent to the equation $y = 3x$,” Empson & Levi, 2011, p. 134). Further, fraction magnitude knowledge appears to fully mediate the relation between early whole number magnitude knowledge and later fraction arithmetic (Bailey, Siegler, & Geary, 2014). That is, children who are highly accurate in their number line estimations of whole numbers in first grade tend to have highly accurate number line estimations of fractions in middle school which then predicts higher fraction arithmetic skills.

Number lines are a visual, mathematically correct way to represent complex fraction concepts. The importance of the number line has recently been noted in the Common Core State Standards for Mathematics (CCSS-M) with students being asked as early as Grade 3 to understand a fraction as a number on a number line (National Governors Association Center for Best Practices & Council of Chief State School Officers [NGACBP & CCSSO], 2010). Unfortunately, many interventions used in schools still emphasize part–whole models for struggling learners, often to the exclusion of number line models (Jordan, Resnick, Rodrigues, Hansen, & Dyson, 2017). Part–whole models represent fractional values as a shaded region of a whole or a subset of a group of objects. For example, a popular part–whole model used when teaching fractions is in the form of a pizza relating slices to fractional values. These models are concrete representations that build on intuitive understandings (Mix, Levine, & Huttenlocher, 1999). However, overemphasis of this approach may lead to a limited way of thinking about fractions that does not encourage an understanding of fractions as numbers with their own numerical magnitude (Gersten et al., 2017). In other words, fractions are numerical values determined by the relationship between the numerator and denominator and not just “parts of a whole.” Although number lines can be thought of more broadly as part–whole models when partitioning, the number line more naturally lends itself to a continuous and unified representation of numerical fraction magnitude than traditional part–whole models. The number line representation is more aligned with the understanding that the numerator and denominator work together to determine the value of the fraction and cannot be separated, as opposed to other tools used to find a discrete part of a whole (e.g., pizza slices). Hamdan and Gunderson (2017) found that even second and third graders trained on a fraction number line estimation task demonstrated greater transfer of knowledge to a novel fraction comparison task than those who were trained using an area model. Previous work with fourth graders deemed at risk for mathematics difficulties showed that the number line is an effective tool for building fraction concepts and skills (Fuchs et al., 2013, 2014), although the extent to which the findings hold over time was not investigated.

Principles From the Science of Learning

General learning principles should guide any mathematical intervention. Findings from the science of learning have yielded many such principles that instructional designers can use in learning environments (Booth et al., 2017). Many factors related to instructional technique, dosage, and timing can be combined in an unlimited number of ways to yield varying effects (Koedinger, Booth, & Klahr, 2013). In the present study, we chose to employ several major principles that improve understanding and retention for students with learning difficulties and disabilities. These include use of explicit instruction, representational gestures, dual coding (i.e., words and visuals) for presenting information, spaced and interleaved practice, and systematic feedback. The following is a description of the learning principles that we capitalize on to increase the effectiveness of our experimental intervention.

Explicit Instruction

Hanover Research (2014) conducted a curricular review of seven mathematics intervention programs that have research support in an attempt to narrow in on key components of effective mathematics interventions. The one common instructional component of these interventions was explicit and systematic instructional methods. In a meta-analysis on mathematics instruction, Gersten and colleagues (2009) also found that explicit instruction was a highly effective instructional technique for students with mathematics difficulties. Explicit instruction encompasses a wide array of instructional approaches but, as explained by Gersten and colleagues, involves teachers demonstrating a step-by-step strategy to solve a specific type of problem and encouraging students to use this particular strategy in their own work. To be effective, explicit instruction must include clear teacher models of problem-solving, opportunities for guided practice, and regular feedback (Doabler & Fien, 2013). Bryant and colleagues (2008) found that an intervention using explicit instruction focused on number, operations, and quantitative reasoning improved mathematics achievement. The instructor in Bryant and colleagues' intervention explained and modeled the steps needed to solve problems before asking students to complete problems on their own. For example, when teaching addition and subtraction involving doubles, the instructor said, "I have 12 connected cubes. I break them into two equal parts. Count with me how many in each (6)" as she made two rows of six cubes. The instructor then said, "This is a double fact: 6 + 6 = 12." After modeling the strategy on the board, students were then given the same materials and asked to model a similar double fact (e.g., 4 + 4 = 8) while the instructor continued to ask guiding questions and was available for further support.

Representational Gestures and Physical Movements

When people speak, they naturally gesture or move their hands to help convey an idea to listeners. Representational gestures depict a spatial object, event, or abstract concept (Goldin-Meadow, 2011). For example, a teacher can indicate that two sides of an equation are equal by using a sweeping hand motion under each side of the equation while also explaining the equivalence of the two sides out loud (Perry, Church, & Goldin-Meadow, 1988). These kinds of gestures enhance learning for mathematical material (e.g., Church, Ayman-Nolley, & Mahootian, 2004; Cook & Goldin-Meadow, 2006; Ping & Goldin-Meadow, 2008). On a Piagetian conservation task, a gestural demonstration of the width of a glass with C-shaped hands placed side by side to approximate the corresponding width improved children's performance more than verbal explanation alone (Ping & Goldin-Meadow, 2008). Ping and Goldin-Meadow argue that these kinds of gestures during instruction help students form abstract representations of the problem to be solved. Although gesturing may not be necessary in mathematics problem-solving situations that involve concrete objects, it becomes particularly useful when students must go beyond whole number arithmetic to working with abstract representations, such as fractions. Carefully designed gestures and movements that emphasize target fraction concepts (e.g., focusing on the magnitude of the fraction by using a range of hand and finger movements to convey fractions of varying sizes) assist students in forming a representation of the concept and focus attention on key visuals.

Dual Coding (Visual and Verbal) and Triple Coding (Magnitude)

Dual coding purports that information is encoded into long-term memory via two pathways: visual and verbal. These pathways are physiologically interconnected but also function independently (Paivio, 1986; Welcome, Paivio, McRae, & Joanisse, 2011). Evidence suggests that retention of material is greatest when visual information is presented in combination with verbal information (Cuevas, 2016, for a review). For example, research comparing the retention of abstract words (e.g., justice) with concrete words (e.g., hammer) reveals that verbal information that is associated with images (i.e., concrete words) is more easily remembered than abstract terms not easily linked to imagery (Bauch & Otten, 2012; Welcome et al., 2011). This finding suggests that pairing verbal concepts with visual cues should be particularly helpful in improving retention of that material. A related model more specific to *numerical* cognition is Dehaene's Triple Code model (Dehaene, 1992), which suggests that numbers are stored in three distinct but related forms. Two of these forms are similar to dual processing: visual form (i.e., Arabic numeral) and auditory form (i.e., verbal word/name). The additional form is that of an analog magnitude representation akin to a mental number line. Our intervention approach supports both of these models.

Spaced and Interleaved Practice

Practice is a critical component of any intervention aimed at skill retention. However, the way in which practice is used affects learning and retention. Practice sessions should be spaced out or distributed over time, rather than practicing only after the corresponding instructional unit. Distributing practice across multiple lessons is more effective for producing long-term retention than practicing in quick succession or in a single session (see Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013, for a review). Although distributed practice effects seem to be greatest for simpler materials in free-recall tasks (e.g., multiplication facts; Don- ovan & Radosevich, 1999; Rea & Modigliani, 1985), there also is evidence demonstrating the benefits of distributed practice for learning more complex mathematical concepts, such as determining simple permutations (Rohrer & Taylor, 2006, 2007).

Interleaved or layered practice involves mixing the types of problems used during practice sessions rather than grouping them according to problem type (Mayfield & Chase, 2002). For example, intermediate-grade students who completed practice assignments with different types of problems mixed together were more accurate in their mathematics problem-solving than students who were assigned blocked practice (i.e., problem sets presented by problem type) over the same period of time (Rohrer, Dedrick, & Stershic, 2015). Interleaved practice helps students evaluate the demands of types of problems and choose the correct strategy according to problem type (Taylor & Rohrer, 2010). Additionally, interleaving practice problems across instructional units increases spacing of the practice (Carpenter, Cepeda, Rohrer, Kang, & Pashler, 2012).

Feedback

Feedback is particularly useful for children with low prior knowledge (Fyfe, Rittle-Johnson, & DeCaro, 2012). However, the timing and manner of the feedback must be considered (Shute, 2008, for a review). Low-achievers in particular benefit from feedback that is immediate (Mason & Bruning, 2001), explicit, and directive (Moreno, 2004). Low-performing students also benefit from feedback that is structured and scaffolded; feedback that is immediate and specific to the incorrect step taken in real time is particularly effective (Graesser, McNamara, & VanLehn, 2005). For example, Fyfe and Rittle-Johnson (2016) found that children with low prior knowledge who incorrectly responded to a mathematics equivalence problem benefitted more from receiving immediate feedback including the correct answer compared with no feedback or summative feedback presented at a delay once all problems were solved.

Intervention Research on Fractions

The majority of published fraction intervention studies focus on fraction arithmetic and solving word problems that require arithmetic skill. For example, Shin and Bryant (2017) conducted a small-scale case study on the effects of a computer-assisted pro gram (*Fun Fractions*) that included metacognitive problem-solving strategies (i.e., teaching a heuristic they term *Read-*

Restate-Represent-Answer) on three students' solving of word problems based primarily on fraction arithmetic. Bottge and colleagues (Bottge et al., 2014) examined the effects of another computer-assisted program (*Fractions at Work*) that centered instruction around video-based real-world problems (e.g., building a skateboard ramp) to be solved by fictitious students within the video along with hands-on applied projects. Though this program used a range of virtual and concrete representations of fractions (e.g., fraction strips, number lines), the primary outcomes of interest were fraction arithmetic.

One experimentally studied fraction intervention focused on students' understanding of fraction magnitudes as a continuous numerical quantity is *Fraction Face-Off!* by Fuchs and colleagues (Fuchs, Schumacher, Malone, & Fuchs, 2011). Fuchs and colleagues (see Fuchs, Malone, Schumacher, Namkung, & Wang, 2017 for a review) have iteratively designed an effective fraction intervention with large effect sizes at immediate posttest for fourth-grade students at risk for mathematics difficulty. Their fraction intervention is supported by domain-general cognitive learning principles (e.g., schema-based instruction, supported self-explaining). However, *Fraction Face-Off!* is designed for younger students when fraction instruction typically begins in schools. Moreover, the durability of the results was not evaluated after a delay. Our intervention builds on promising findings by Fuchs and colleagues through targeting sixth-graders who have already received fraction instruction yet still struggle with fraction understandings and assessing effects over time.

The Present Study

The present randomized study evaluated the effectiveness of an experimental intervention designed to build fraction knowledge in students who reach sixth grade with low fraction skills. As noted, the experimental intervention focused on the number line, although other common representations of fractions were also introduced (e.g., area models) and connected explicitly to the number line. The experimental intervention also incorporated the key learning principles described previously. It was carried out at the beginning of the school year to help students benefit from their regular fractions instruction, which was occurring concurrently. Students were selected for inclusion based on low performance on a validated fraction screener (Rodrigues, Jordan, Hansen, Resnick, & Ye, 2017). Relative to a business-as-usual control, we assessed learning not only from pre- to immediate posttest but also at a 7-week delay. The delayed posttest is particularly important for determining skill retention over time (e.g., Bailey et al., 2016). Outcomes included assessments that measured specific fraction skills of fraction number line estimation (FNLE), fraction magnitude comparisons, and fraction arithmetic as well as a broader fraction concepts measure. Skills measured on the broad fraction concepts measure as well as the FNLE task and fraction comparison task were all directly targeted within intervention lessons. In addition to expected improvements on the broad concepts and magnitude measures, we were interested in the extent to which students in the experimental intervention group would show greater improvements than control group students on fraction arithmetic skill, which was not the explicit focus of the experimental intervention.

We also assessed children’s general cognitive competencies. Prior work suggests that fadeout effects (i.e., the finding that mathematics intervention treatment effects often diminish over time once the intervention has been complete) may be attributable to preexisting differences (e.g., language; prior mathematics knowledge) between students in treatment and control groups (Bailey et al., 2016; Bailey, Fuchs, Gilbert, Geary, & Fuchs, 2018). Although not a major focus of the current study, assessing students on a range of cognitive measures enabled us to test and account for differences in these competencies in our analyses. That is, we collected these measures with the intention of using measures that conditions differed upon as covariates. We assessed working memory, receptive vocabulary, inhibitory control, nonverbal matrix reasoning, and nonsymbolic proportional reasoning, all known predictors of mathematics achievement (e.g., Fuchs et al., 2014; Hansen et al., 2015; Ye et al., 2016). We also assessed classroom attention. Struggling students commonly display poor attention (Fuchs et al., 2005), and students' classroom attentive behavior has been shown to influence fraction knowledge in particular (Hansen et al., 2015; Hecht, Close, & Santisi, 2003; Resnick et al., 2016; Ye et al., 2016). Demographic variables including gender, special education status, and English language learner status were assessed to ensure equivalence of conditions.

In sum, the current study evaluated the effectiveness of an intervention centered primarily on the number line, one that also incorporates cognitive learning strategies to improve learning and retention. In addition to a broad measure of fraction concepts more akin to school measures, we were concerned with the specific conceptual skills of fraction magnitude estimation for sixth grade students at risk for mathematics failure. We hypothesized that the intervention would lead to meaningful immediate and longer-term improvements on a range of fraction skills. We also explored whether the experimental intervention would have differential impacts on the learning of students with different cognitive or behavioral competencies, focusing on classroom attentive behavior in particular.

Method

Participants

Students from two public middle schools in the Northeast region of the United States were recruited to participate. Both schools were racially and ethnically diverse (School 1: 36% Black non- Hispanic, 32% white Hispanic, 12% white non-Hispanic, 12% Other; School 2: 50% Black, non-Hispanic, 31% white Hispanic, 19% white non-Hispanic) and served students from low-income families (School 1: 52% low income; School 2: 35% low income). Qualifying for enrollment in the schools' free or reduced lunch program was used as a proxy for low-income status. Information on SES was not available at the student level. In the previous school year, proficiency level on the state test in mathematics was 45% for both schools. That is, the majority of students did *not* meet state mathematics proficiency benchmarks.

Using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), we conducted a priori power analyses to determine the minimum sample size to detect a meaningful effect on fraction knowledge using $\alpha = .05$. Our preliminary work led us to expect large effects (i.e., Hedges' g = 0.8) of the experimental intervention on the fraction concepts measures. A priori analyses suggested a sample size of 54 would provide power of .81 to detect a large effect on fraction concepts.

A fraction screener (described in the Method section) was administered to all nonhonors sixth graders in both schools. Students who scored at or below a validated cutpoint for mathematics difficulties were invited to participate in the study (Rodrigues et al., 2017). Ninety-nine of the 392 students (25%) screened met this selection criteria and were invited to participate in the experiment. Of this sample, parents of 56 students provided informed consent (School 1: n = 27; School 2: n = 29). These students also assented to participate.

Participants were randomly assigned either to the experimental intervention condition or the business-as-usual control (BAU) condition stratified by regular mathematics classroom. If there were three participating students in a classroom, we assigned the extra student to the experimental intervention group to ensure the target intervention group size of three to four students, an optimal group size for struggling learners (Fuchs et al., 2014). Participants within the experimental intervention condition were then randomly assigned to one of four experimental intervention groups within each of the schools (i.e., eight experimental intervention groups total), disregarding mathematics classroom membership to reduce issues of nesting (i.e., each group included students from several mathematics classes). Each small group was taught by one researcher- instructor. Because of available resources, four researcher– instructors taught one group each and two researcher–instructors taught two groups (one in each school).

Of this original sample, three students in the experimental intervention and one student in the control group moved away before the experiment was completed; one experimental intervention student was removed before the completion of the study because of overly disruptive behavior. The final sample (*N* = 51; 20 males, 31 females) included students enrolled in seven different classrooms across the two schools (School 1: *n* = 25; School 2: *n* = 26). Five participants were English Language Learners (ELL; School 1: *n* = 4; School 2: *n* = 1). There were 28 experimental intervention students and 23 control group students.

Pretest and Posttest Fraction Measures

Fraction concepts. This broad fraction concepts measure was made up of 24 released National Assessment of Educational Progress (NAEP) items that assessed various aspects of fraction concepts. Nineteen of these items were used as a screener to select students eligible for participation. The screener included six part– whole area model items, one set model item, six equivalence items, two fraction magnitude items, one estimation item, and three comparison and ordering items. These 19 screening items were validated in a prior study (Rodrigues et al., 2017), which used receiver operating characteristic (ROC) analysis to assess the diagnostic accuracy of the screener for predicting students' later mathematics risk status. The fraction concepts screener yielded an area under the curve (AUC) value of .881, signifying a very good screener

(Cummings & Smolkowski, 2015). The analysis identified that students who scored at a cut score of 10 or below on the 19-item screener had an 87% chance of failing a standardized mathematics test at the end of sixth grade.

Five additional, more challenging, NAEP items were added to the fraction concepts measure to avoid potential ceiling effects and were administered along with the 19 screener items at all time points. The five items assessed part–whole understanding of an area model, fraction equivalence, and fraction magnitudes. As these items were not included in the prior screener validation study, performance on these five items was *not* counted toward the screener score. Students received one point for each correct item for up to a total of 24 points for the entire measure (up to a total of 19 points for the screener). Screener scores were used only for inclusion criteria. Performance on the full 24-item measure was used at all time points to assess the effectiveness of the intervention. The full fraction concepts measure as well as the screener items have displayed good internal consistency with sixth graders in prior studies with larger samples from the same region (a= .86 and .78, respectively; Jordan et al., 2017; Ye et al., 2016). Sample NAEP items are displayed in Figure 1.

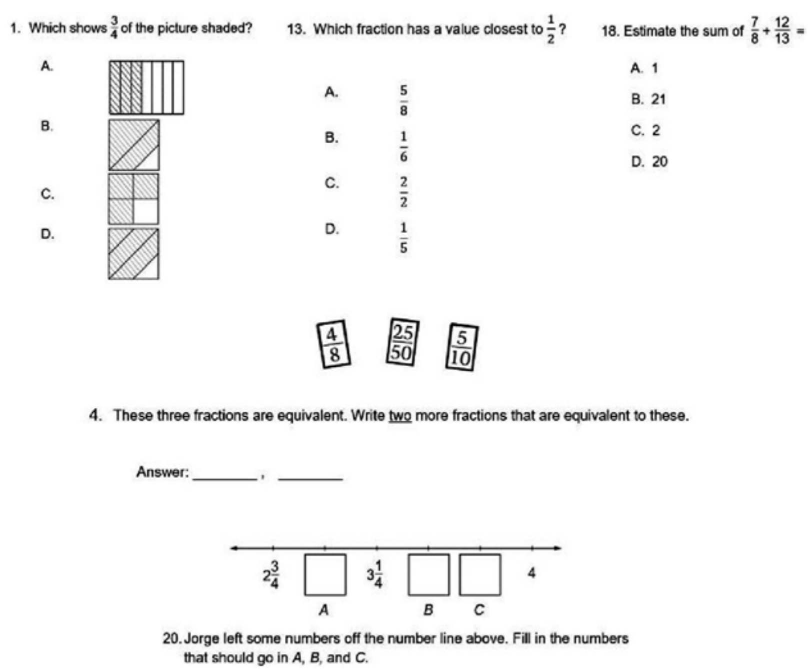

Figure 1. Sample of released NAEP fraction conceptual items.

Fraction comparisons. The paper and pencil fraction comparisons measure includes 24 items for which students are presented with two fractions and asked to select the larger fraction. Students complete as many problems as they can within three minutes. The measure includes several types of comparisons including unit fractions (e.g., 1/3 or 1/2), fractions with like denominators (e.g., 5/7 or 6/7), fractions with like numerators (e.g., 2/4 or 2/5), reciprocal fractions (e.g., 8/4 or 4/8), and fractions with different denominators and numerators (e.g., 12/50 or 8/60). Students received one point for each correct item for up to a total of 24 points. Internal reliability of the fractions comparison measure in sixth grade with a larger sample from the same region was high (α =.92; Jordan et al., 2017).

Fraction number line estimation. Two paper-and-pencil fraction number line estimation (FNLE) tasks were used. The tasks were adapted from a computer administered version (Hansen et al., 2015) for logistical reasons. These tasks includeda0 to1 scale and a 0 to 2 scale. The 0 to 1 scale included six individual number lines and the 0 to 2 number line included eight individual number lines, each 81 mm in length. The number 0 was placed below the left end of the number line and the 1 or 2 was placed at the right end of the number line. The target fraction was centered below each number line. The number lines for each task were printed on the same sheet of paper but were staggered so that participants could not easily use their estimates on other number lines to inform their placements. Students were instructed to “Mark a line to show where each number belongs on the number line if the endpoints are 0 and 1 [or 2].” The target numbers on the 0 to 1 task were all proper fractions (1/4, 1/5, 1/3, 1/2, 1/19, 5/6). The target numbers on the 0 to 2 task included proper fractions (3/8, 5/6, 1/2, 1/19), improper fractions (7/4, 5/5), and mixed numbers (1 11/12, 1 1/2).

Prior work on *whole number* magnitude has been concerned with the logarithmic to linear shift that occurs throughout development with increasingly large scales. That is, whole number magnitudes are represented in a compressed logarithmic distribution (i.e., overestimating the distance between smaller numbers and underestimating the distance between larger values) and become increasingly more linear throughout development with increasingly larger scales (Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Booth & Siegler, 2006; Siegler & Booth, 2004; Thompson & Opfer, 2010). This method requires fitting linear and logarithmic functions for each participant's estimates and then determining which pattern is a better fit for that learner (Siegler & Opfer, 2003). Whereas becoming increasingly linear in one's whole number estimates is the standard developmental trajectory that children follow (leading to increases in accuracy with whole numbers), the same is not true for fractions (e.g., Opfer & DeVries, 2008; Siegler, Thompson, & Opfer, 2009). Thus, most work on *fraction* magnitude estimation uses percent absolute error (PAE) as a performance measure on fraction number line tasks (e.g., Fazio, Kennedy, & Siegler, 2016; Fuchs et al., 2017; Hamdan & Gunder- son, 2017; Siegler & Pyke, 2013). The current study follows this field standard.

To gauge students' accuracy of their estimates, PAE was calculated for each estimate by dividing the absolute difference between the participant's estimate and the accurate target location by the scale of the estimates. The PAE for each of the 14 number lines was averaged to determine the overall PAE. Lower scores indicate more accurate performance. All number line estimates were measured twice by the same two research assistants, and any discrepancies were reconciled. Internal consistency for the fraction number line task in sixth grade with a larger sample from the same region was high (α = .87; Resnick et al., 2016).

Fraction arithmetic. There were 12 written fraction arithmetic items: four addition, five subtraction, and three multiplication. Proper and improper fractions as well as mixed numbers were used. Addition and subtraction items included addends and subtrahends both with like and unlike denominators (e.g., 3/4 + 2/3 = ___; 5/6 – 2/6 =___). In each of the three multiplication items, one factor was a fraction and one factor was an integer (e.g., 3/4 X 12 =_). Students were asked to give their answers in simplest form. Students received one point per correct item as well as one point per correct response in simplest form. Thus, there were a total of 24 points to be earned. The internal reliability for this measure in sixth grade with a larger sample from the same region was high (α = .82; Hansen et al., 2015).

General Competencies

We assessed working memory, receptive vocabulary, and inhibitory control with validated measures from the NIH (National Institutes of Health) Toolbox (Gershon et al., 2013). All toolbox tasks were administered individually on an iPad. All raw scores were converted to scaled scores for ease of interpretation (*M* = 100, *SD* = 15). These three NIH toolbox measures have demonstrated good to excellent test/retest reliability at 8–15 years of age, with alphas ranging from .81–.91. Convergent and discriminant validity were also established (see Bauer & Zelazo, 2013).

Working memory. Participants were presented with pictures of food and/or animals that were labeled with audio and text and displayed for 2 s. Participants recalled items aloud from smallest to largest, with the number of images increasing by one item per trial. Students completed a unidimensional (either food or animals) and a two-dimensional (both food and animals) list. Higher scores represent more items correctly recalled from both lists.

Receptive vocabulary. Participants were asked to select the correct image from a group of four that most closely matched the meaning of the word presented. The Toolbox Picture Vocabulary Test (TPVT) is a Computer Adaptive Test (CAT) that adjusts the level of difficulty based on each student's performance. Item Response Theory (IRT) is used to score performance on the TPVT. Higher theta scores represent better vocabulary.

Inhibitory control. Students completed a traditional flanker task in which they were shown a series of arrows and asked to choose the direction of the center arrow. Sometimes the direction of the center arrow was congruent with the arrows flanking it and sometimes the direction was incongruent. Higher scores represented greater speed and accuracy.

Nonverbal reasoning. Nonverbal reasoning was assessed using the Matrix Reasoning subtest of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). Students were shown a series of grids that contain pictures in three of the four cells which begin a pattern and are asked to choose the next grid that completes the pattern. Higher *t* scores (M = 50, SD = 10) represent more items correct. Test/retest reliability was adequate at 12–16 years of age (α = .74). Convergent, discriminant, and construct validity were also established (see Wechsler, 1999).

Nonsymbolic proportional reasoning. Proportional reasoning was measured using an iPad adaptation of the nonsymbolic scaling task used by Boyer and Levine (2012). Students are shown a target “juice” mixture represented by a vertical bar with a portion of red representing the red powder mix and a portion of blue representing the water. Students are shown two additional bars that are a different size than the target and tasked with choosing the option that is a rescaled version of the original. Correct responses required either scaling down from a larger target to a smaller match or scaling up from a smaller target to a larger match. Higher scores represent more correct trials. Internal reliability with a larger sample from the same region was high (α = .93; Ye et al., 2016).

Classroom attentive behavior. Classroom attention was measured using mathematics teacher reports on the Inattentive Behavior subscale of the SWAN Rating scale (Swanson et al., 2006). *Attentive behavior* is one’s level of attention exhibited in the classroom as observed by a student’s regular classroom teacher. This subscale included nine Likert-style items that followed *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) criteria for attention-deficit/hyperactivity disorder. A sample item is “Gives close attention to detail and avoids careless mistakes.” Responses ranged from 1 (*far below average*) to 7 (*far above average*). Lower scores represent poorer attention. Sum scores range between 9 and 63. Students’ regular mathematics teachers rated individual students’ behavior. Internal consistency is high (α = .92; Lakes, Swanson, & Riggs, 2012).

Background Variables

Demographic variables including gender, special education status and English Language Learner (ELL) status as well as whether students were classified as having a math, reading, and/or behavioral disability, were obtained through school records with permission from parents/caregivers. Income status, race, and ethnicity were not available for individual students.

Procedure

The study design included group administration of a fraction pretest right before the intervention period, a posttest immediately after the intervention period, and a delayed posttest seven weeks later. The general cognitive measures were administered prior to the intervention period with the exception of the SWAN attentive behavior scale (Swanson et al., 2006).
As the intervention was administered at the start of the school year, the SWAN Rating Scale was completed several weeks after the intervention began to allow general classroom teachers to become acquainted with their new students so that they could provide more accurate representations of their normal classroom behavior. Posttest fraction measures were administered to the experimental and control conditions at the same time in the same classroom.

The intervention took place during a 6-week period in which all students received specialized help from a teacher within their school. This designated 45-min intervention time is *in addition to* students' regular mathematics class. In their regular mathematics classes, both schools used the same mathematics curriculum: *Connected Mathematics Project* (CMP; Lappan, Difanis Phillips, Fey, & Friel, 2014). The CMP curriculum is aligned with CCSS-M (NGACBP & CCSSO, 2010). According to its manual, it aims to help students develop mathematical understanding by emphasizing connections between mathematical ideas and their real-world problem-solving applications. When followed according to its intended design, CMP includes minimal explicit instruction on specific strategy use but instead encourages students to invent their own strategies for problem-solving and discuss multiple strategies during whole group discussions. For fractions in sixth grade, this curriculum covers factors, models, and fraction operations.

During the additional class period dedicated to intervention, students in the experimental intervention condition received 27 researcher-designed lessons (described further below). These lessons were administered to each of the small groups by one of the trained instructors. Concurrently, students in the control condition received their regular mathematics intervention provided by their school. Both schools used a computer adaptive tutoring software for their mathematics intervention period, on which students worked individually. Students received individualized assistance in mathematics from the computer adaptive software based on their current level of performance. One school used *Dreambox Learning* (2012). The other used *i-Ready* (2016). Both programs are aligned with the CCSS-M and as such address fraction understanding including operations and some magnitude judgments along with other sixth-grade mathematics content.

Instructor Training

Experimental intervention instructors were trained research assistants who also participated in lesson design. Instructors varied in prior teaching experience. Two instructors were doctoral students, two were postdoctoral researchers, and two were previous certified teachers. Each of the six instructors received more than 16 hr of the same training in administration of the lessons from one of the authors of the current paper. Training included practice in use of gestures, proper strategies for providing feedback, instructor/student dialogue, and behavior management. Experimental intervention instructors also practiced teaching the lessons in pairs and provided each other feedback for lesson improvement prior to administering the lessons.

Experimental Intervention Design

Lesson structure. The lessons were situated in the context of a color run—a race in which runners are showered with different colored powders at stations along the way. The race course models a number line, on which students can think about fractions and their magnitudes in a real-world context for understanding fraction magnitude (Rodrigues, Dyson, Hansen, & Jordan, 2016). Lessons focused primarily on denominators that occur frequently in measurement activities of daily life, including halves, thirds, fourths, sixths, eights, and twelfths. However, students also practiced with other denominators during practice activities and games involving fraction magnitude comparison. The lessons were carefully scripted to increase fidelity. The scope and sequence of the experimental intervention lessons is presented in Table 1. A general overview of the lesson structure is presented in Table 2.

Prior to the explicit instructional time of each lesson, students completed a warm-up worksheet in which they individually practiced material they learned from the previous lessons. The National Mathematics Advisory Panel (NMAP, 2008) deems the ability to recall basic mathematics facts a crucial prerequisite skill for mathematics success. Quick and accurate recall, or automaticity, of basic mathematics facts is thought to free up the necessary attentional resources that a learner needs to focus on more complex aspects of a task. Students who have a learning disability and those with low mathematics achievement typically struggle with both accuracy and speed in mathematics fact recall (Geary, 2004; Jordan, Hanich, & Kaplan, 2003). Therefore, after the warm-up, students participated in practice of whole number multiplication facts which targeted multiplicands used as the main denominator of that day's lesson. Practice during a given lesson included facts learned in prior lessons and was interleaved for fact retention.

Next, as a group, students practiced counting aloud fractions of like denominators along the number line using both proper and improper fractions or whole and mixed numbers. These activities were also aimed at automaticity in preparation for the explicit instructional period focused on concepts which lasted about 15 to 20 min per lesson. These lessons were predominantly focused on developing students' understanding of fraction as a number or magnitude on their mental number line. As such, this was accomplished mostly through the use of a visual number line (or "race course") but also through the use of fraction strips and other representations. Although not the primary focus, some attention was given to addition and subtraction in a conceptual manner as represented by moving forward (addition) and backward (subtraction) along a number line. More details are presented in Table 1, which displays the scope and sequence of the lessons provided within the experimental intervention.

Each lesson concluded with short, fast-paced card games that gave students opportunities to practice lesson goals related to fraction concepts in particular. The games targeted fraction concepts such as magnitude judgments and their corresponding strategies (e.g., comparing two fractions to each other, to one half, etc.) as well as fraction equivalencies (e.g., 3 is the same as how many halves?; 3 halves is the same as how many fourths?). Lastly, students completed an independent cool-down worksheet in which they solved problems that assessed their knowledge of the concepts presented in the instructional period as well as concepts from previous lessons. Many of the practice problems were *near* transfer problems that required students to apply previously learned skills to find solutions to problems very similar to those practiced throughout the lessons. However, *far* transfer (i.e., novel) problems were also included to encourage students to expand upon their knowledge and modify methods learned to find a solution (Barnett & Ceci, 2002).

Cognitive learning strategies. Strategies based on key learning principles were implemented comprehensively throughout the lessons to target the range of skills covered within the lessons as noted within the Scope and Sequence in Table 1. For example, students completed spaced and distributed practice on numerous skills including but not limited to fraction magnitude comparisons, whole number multiplication facts, and partitioning and marking number lines. The following sections provide examples of how each principle was regularly implemented.

Explicit instruction. In each experimental intervention lesson, students were explicitly taught concepts and step-by-step strategies the majority of the time (i.e., 15–20 min per lesson) prior to practicing them independently. This explicit instruction period focused primarily on the number line representation of fractions although complementary representations were also provided (e.g., fraction bars). For example, in Lesson 2 students were taught how to partition a number line race course into half miles. The instructor displayed a blank number line that represents a four-mile race course and then used a paper bar that represented one mile to mark off each of the four miles. Once all students correctly partitioned and marked their number lines into four miles, the instructor demonstrated finding one half mile by finding the midpoint of the first mile. The instructor placed a pencil under the line and estimated at which point the segments on the left and right of the pencil are equal. The instructor then made a mark at this point and explained that this mark is the one-half mile mark, because it is one of two equal portions of a whole. The instructor repeated this process until the remaining halves were marked, ensuring that students followed along on their own race course. This explicit modeling was included to encourage students to adopt these efficient partitioning strategies in their independent work. A similar process was used for fractions with other denominators such as fourths and eighths (see Figure 2).

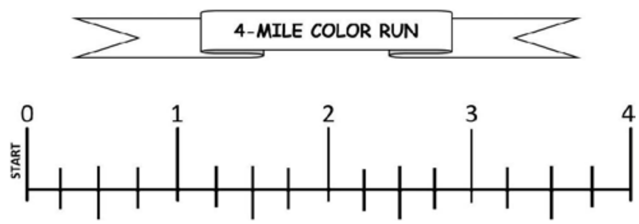

Figure 2. *Example of number line used in lesson on fourths.*

Representational gestures. In each experimental intervention lesson, gestures were used to represent corresponding concepts and ideas. A *chunking* gesture helped students visualize the magnitude of a particular unit fraction along a number line. The instructor placed her index finger and thumb at the start and endpoint of a particular magnitude (e.g., 0 to 1/4) much like a bracket and held her fingers in this manner while demonstrating moving along the number line in consistent units as she counted aloud and explained the meaning of the corresponding magnitude. This gesturing was done to foster the understanding that the magnitude of any given fraction on a number line is the distance between 0 and that fraction—not simply an arbitrary label for the mark on the number line. Other gestures highlighted the meaning of the numerator and denominator. For example, when explaining the meaning of one half, the instructor displayed the fraction on the board and used a *components* gesture. This gesture involved pointing to the “1” with one finger while saying “one of . . .” and pointing to the “2” with two fingers while stating, “two equal portions in the whole.”

Dual coding. Verbal information was typically presented with a visual aid. Along with the number line, the experimental intervention used concrete materials such as fraction bars and other magnetic manipulatives that were presented and used with corresponding verbal explanations. For example, when discussing equivalent fractions, the instructor presented students with magnetic fraction bars that displayed the same whole partitioned into different numbers of parts and labeled accordingly (e.g., halves, fourths, and eighths). When converting halves to fourths, the instructor connected verbal explanations with visual representations by saying, “We know that each green magnet represents one-half. Do you see an equivalent fraction for one-half? One-half equals two-fourths. One-half and two-fourths are equivalent fractions.” Students were encouraged to match up the one half magnet with two one fourth magnets to connect these verbal explanations with visual representations. An image of these magnets is displayed in Figure 3.

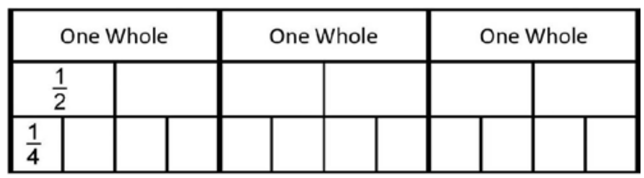

Figure 3. Fraction magnet manipulatives.

Spaced and interleaved practice. Distributed and interleaved practice was used to enhance long-term retention of basic prerequisite competencies (e.g., whole number multiplication facts) and develop fluency in specific problem-solving strategies for more conceptual tasks (e.g., fraction magnitude comparisons). For example, students practiced whole number multiplication facts for several minutes during each lesson over the course of the 27-day experimental intervention. Students also practiced fraction magnitude comparisons during which they were shown two fractions and asked to indicate which fraction is more. Fractions taught in previous lessons (e.g., unit fractions) were interleaved with those taught in later lessons (e.g., improper fractions). Students also worked on earlier presented problem types both at the warm-up and cool-down phases of the lessons. For example, during the warm-up and cool-down activities within a lesson, students practiced partitioning a number line using a range of denominators and completing fraction computations with fractions of varying denominators, even though the bulk of each lesson was generally structured around one specific denominator that varied on every fifth lesson (e.g., Lessons 1– 4 focused on halves, Lessons 5– 8 focused on fourths).

Feedback. Instructors gave individual feedback on students' independent work during the warm-up activity and throughout the lessons. This feedback was both corrective and process-oriented, allowing students to revise their strategies. One example of *corrective* feedback was the process used for error correction while practicing whole number multiplication facts. When a student gave an incorrect product, the instructor briefly displayed the number sentence with the correct product and then gave the student another opportunity to answer the question. An example of *process-oriented* feedback was the approach used for addressing incorrect fraction comparison strategies. During practice on fraction comparison tasks, students were asked to note the best strategy to use for each particular pair of fractions based on fraction type. For example, when comparing 6/8 and 1/4, if a student named an incorrect strategy (e.g., "more parts equals larger fraction"), the instructor noted the best strategy (e.g., benchmarking – determine which is closer to 1) and explained why this choice was more efficient (e.g., "These fractions have different denominators so we can't simply compare number of parts. We know that 1/4 is closer to 0 and 6/8 is closer to 1. Therefore, 6/8 is larger than 1/4"). Common errors across student work (e.g., skipping the fractions that are equivalent to whole numbers when labeling halves on a number line) were discussed and corrected with the group during lessons.

Fidelity of Implementation

Lessons were scripted to ensure instructor fidelity of implementation. All lessons were audio recorded to obtain measures of fidelity, which were transcribed and coded by trained researchers. Each lesson had a checklist of an average of 80 items to check off per lesson (approximately one check per 30 seconds of audio recording). Eight of the 27 lessons were randomly selected for each of the six instructors and coded according to conformity to the scripted activities. Instructors administered an average of 99% of all scripted experimental intervention activities to students. Experimental intervention activities that were not administered mainly resulted from lack of time as a result of natural classroom disturbances (e.g., announcements, fire drills, student tardiness, behavioral issues, etc.).

Plan of Analysis

Evaluating the effects of intervention. The primary focus of the current study was to evaluate the effectiveness of our experimental fraction intervention on improving student performance on a range of fraction skills from pretest to immediate posttest and to determine whether these skills were retained at a delay. To assess whether there were differential effects on fraction learning from pre- to immediate and delayed posttest between experimental and control students, four 2 (Group: Intervention and Control) X 3 (Time: Pretest, Posttest, Delayed posttest) mixed analyses of covariance (ANCOVA) were planned on fraction concepts, FNLE, fraction comparisons, and fraction arithmetic, respectively.

A mixed ANOVA is a general linear model similar to a combination between a repeated measures ANOVA (RMANOVA) and a one-way ANOVA (Murrar & Brauer, 2018). It is termed a mixed ANCOVA once a covariate is added to the model. In the absence of baseline balance, we planned to control for any cognitive or demographic variables that conditions differed upon at pretest. The planned mixed ANCOVAs explain variance in the dependent variable between groups, within groups, and, most importantly, between groups over time, while controlling for any time invariant covariate measured only at pretest. The main focus of a mixed ANOVA is the Time X Condition interaction, which examines whether differences over time in the dependent variable were significantly different between conditions. Planned trend analyses on Time X Condition interactions involved fitting linear and quadratic terms to determine whether the effect of condition on the outcome increased linearly over time and whether the effect leveled off showing a retention at delayed posttest. All statistical analyses were performed using SPSS v. 24.0 (IBM Corp. 2016).

Individual differences exploratory analyses. We explored whether the experimental intervention had differential impacts on the learning of students with different cognitive or behavioral competencies. A preliminary examination of the data suggested that the effect of intervention on the fraction outcome measures at posttest may be moderated by students' attentive behavior (as measured by the SWAN teacher rating scale). Thus, we examined this possibility using ordinary least squares (OLS) regression models to test for an interaction between students' attentive behavior score and condition on each outcome measure that demonstrated differential improvements based on the mixed ANCOVA results. Because a significant amount of learning occurred specifically between pre- and posttest, models were tested on immediate posttest only. A dummy code variable was created for condition (Experimental intervention = 1; Control condition = 0). SWAN scores were centered to reduce the risk of multicollinearity that may arise when including interaction terms in analyses.

An interaction term was created between the centered Attentive Behavior variable and condition dummy code. A series of multiple regressions predicting posttest scores were conducted. Four regressions for each of the outcomes of interest were built starting with a pretest-only model and adding one term to the prior model until the final model included pretest, dummy code for intervention, classroom attentive behavior, and an interaction term between condition and classroom attentive behavior. The main effects models reestablish the relationship between posttest scores and condition, controlling for pretest scores on each corresponding measure. The three moderation models tested whether the effect of condition on each outcome measure varied by students' attentive behavior as represented by classroom teacher ratings on the SWAN. These exploratory models are displayed in Table 10 and supplemental tables S11 and S12 and results are discussed below.

Corrections for multiple tests. Because the current study was proposed to include planned multiple comparisons within the mixed ANCOVAs noted above, corrections are not necessary (Armstrong, 2014; Perneger, 1998). Additionally, exploratory analyses that do not test specific hypotheses but rather provide suggestions for future work do not require corrections. However, we have opted to take a more conservative approach in interpreting our findings in both sets of analyses and have adjusted our alpha levels for multiple tests. We employed Benjamini and Hochberg's (1995) correction procedure for multiple tests which decreases the False Discovery Rate (FDR). The FDR is the expected proportion of the rejected null hypotheses which are incorrectly rejected. Unlike the classic Bonferroni correction (Bonferroni, 1936), which adjusts the alpha level once to use for all comparisons, the BH correction adjusts the alpha level down to an increasingly conservative cutoff, using an ordered set of obtained m p values, only after each statistically significant result and not after nonsignificant results After finding the largest p value that satisfies $p_k \leq \frac{k}{m}\alpha$, all tests with smaller p values are declared significant. BH corrections were applied to the four Time X Condition interactions within the mixed ANCOVAs with adjusted alpha levels of 0.05, 0.0375, 0.025, and 0.0125. In our exploratory analyses, we adopt BH adjusted alpha levels of 0.05, 0.033, and 0.017 to interpret the results of the three Attentive Behavior X Condition interactions.

Nesting as nuisance. We did not have substantive questions on the effects of classroom level variables in the current study. As previously noted, participants were randomly assigned at the student level within classroom to the intervention or control group. Then, students in the intervention condition were randomly assigned to one of four intervention groups in each school. Thus, we did not expect to find substantial classroom effects. Indeed, intraclass correlations (ICC) were low on all outcomes (ICC < .05) and did not warrant multilevel modeling (MLM). Thus, clustering can be considered a nuisance variable (Clarke, Crawford, Steele, & Vignoles, 2010, p. 7). However, we tested fixed effects models controlling for cluster (i.e., accounting for all cluster-level effects) and to reduce the issue of the omitted variable bias (Huang, 2016; Kennedy, 2003). These exploratory analyses revealed a consistent pattern of results regarding the effects of the intervention on all four fraction outcomes demonstrated in the mixed ANCOVAs as well as the significant moderation of effect of condition by Attentive Behavior in the exploratory analyses reported below.

Determining effect sizes. To determine the magnitude of the intervention effects on each outcome, we utilize Hedges' *g*, Cohen's U3, and an improvement index, all presented in Table 9.[1] As suggested by the WWC *Procedures and Standards Handbook, Version 3* (U.S. Department of Education, Institute of Education Sciences, 2014), effect sizes of an education intervention that uses the same measure for pre- and posttests should be calculated as the difference between the pre- and posttest mean difference of the experimental intervention condition and the pre- and posttest mean difference of the control condition. Thus, effect sizes reported in Table 9 represent the *difference of the differences* as opposed to simply the difference

[1] Whereas Hedges' *g* provides a standardized index of a mean difference of learning between the two groups, partial eta squared (' η_p^2) provides a measure of variance explained by each factor of the mixed ANCOVAs. Therefore, η_p^2 is also presented as an additional measure of effect size in the tables for each of the four 3 X 2 mixed ANCOVAs conducted.

between posttests. Hedges' g provides a better estimate of effect sizes in small samples than Cohen's *d* (Cummings, 2012). Hedges' g can be interpreted using Cohen's (1988) standards of small (0.2), medium (0.5), or large (0.8; e.g., Lakens, 2013). However, Cohen suggested caution when using rules of thumb and emphasized the importance of considering effect size in context. According to the same WWC *Procedures and Standards Handbook*, an effect size (Hedges' *g*) of at least .25 is considered meaningful in education research, even if statistical significance is not reached. An effect size of .25 indicates that the experimental group performed one fourth of a standard deviation higher than the control group, based on the pooled variance from the sample. To provide another measure of practical importance, we also converted the effect size of Hedges' *g* to Cohen's U3 to yield an improvement index. The U3 represents the percentile rank of a student in the control group who performed at the level of an average experimental group student. By definition, the average control group student would rank at the 50th percentile. Thus, finding the difference between the computed value of the U3 for the experimental group and 50% yields the improvement index.

Results

Descriptive Statistics

Descriptive statistics for the entire sample as well as by condition are presented in Table 3 on demographic variables as well as participants' scores at pretest on fraction measures and cognitive competencies. All continuous variables were normally distributed. Prior to addressing our research questions, chi square analyses and *t* tests were conducted to confirm equivalence between conditions on demographic variables, teacher report of attentive behavior, pretest fraction measures and cognitive measures. As demonstrated in Table 3, there were no significant differences at pretest, with the exception of receptive vocabulary. Students in the control condition had significantly higher receptive vocabulary scores (*M* = 81.37, *SD* = 6.99) than the experimental intervention (*M* = 77.55, *SD* = 5.60), $t(49) = 2.166$, $p = .035$. Thus, receptive vocabulary was used as a covariate in the four planned mixed ANCOVAs to control for these differences. No other covariates were used.

Effect of Intervention on Fraction Outcomes

Results from the four 2 (Condition: Intervention and Control) X 3 (Time: Pretest, Posttest, Delayed posttest) mixed analyses of covariance (ANCOVA) conducted on fraction concepts, FNLE, fraction comparisons, and fraction arithmetic (controlling for receptive vocabulary) are presented in Table 4. Estimated marginal means reported for each outcome are adjusted for receptive vocabulary at *M* = 79.27. Greenhouse-Geisser corrections were used to correct for violations of assumptions of sphericity. As previously noted, the Time × Condition interactions are interpreted after applying BH corrections (Benjamini & Hochberg, 1995).

Fraction Concepts

As displayed in Table 5, there was no main effect of time, but there was a main effect of condition as well as a significant time by condition interaction. A trend analysis explicating the time by condition interaction revealed both a significant linear and quadratic component. This trend is plotted in Figure 4, which displays a general increase in scores from pretest to immediate posttest and a leveling off between posttest and delay for experimental intervention students. The control group shows a small but noticeable improvement between immediate and delayed posttest. Post hoc pairwise comparisons revealed that the experimental intervention group had higher fraction concepts scores than the control condition both at posttest ($p < .001$) and delayed posttest ($p = .008$).

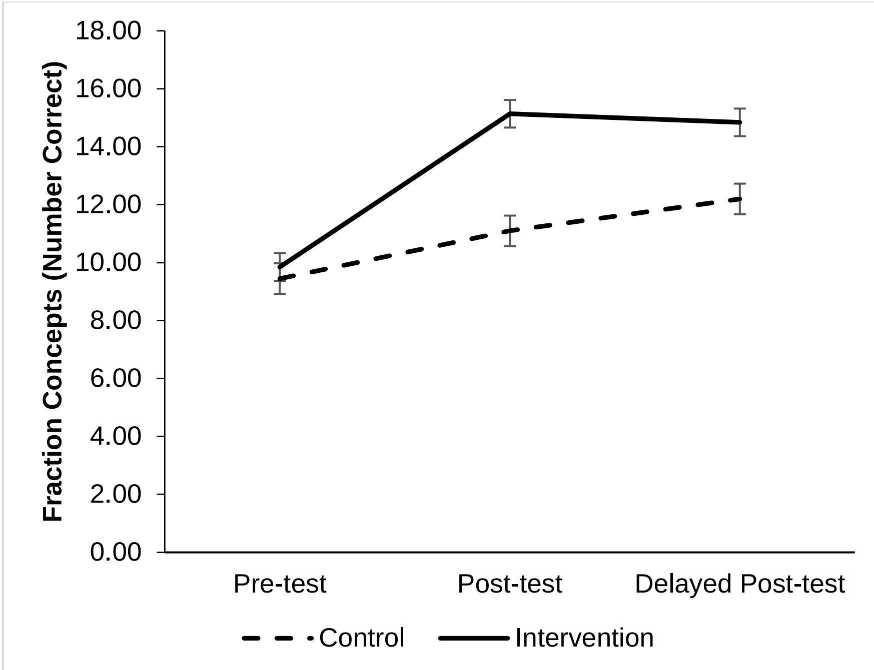

Figure 4. Estimated marginal means of fraction concepts controlling for receptive vocabulary

Fraction Number Line Estimation

As displayed in Table 6, there was no main effect of time, but a main effect of condition as well as a significant time by condition interaction. A trend analysis explicating the time by condition interaction revealed a significant quadratic component and a trend toward a significant linear component. This trend is plotted in Figure 5, which displays a general decrease in PAE (indicating more accurate performance) from pretest to posttest for the experimental intervention group and then a leveling off between posttest and delay. Post hoc pairwise comparisons revealed that the experimental intervention group had lower PAE than the control condition, both at posttest ($p < .001$) and delayed posttest ($p < .001$).

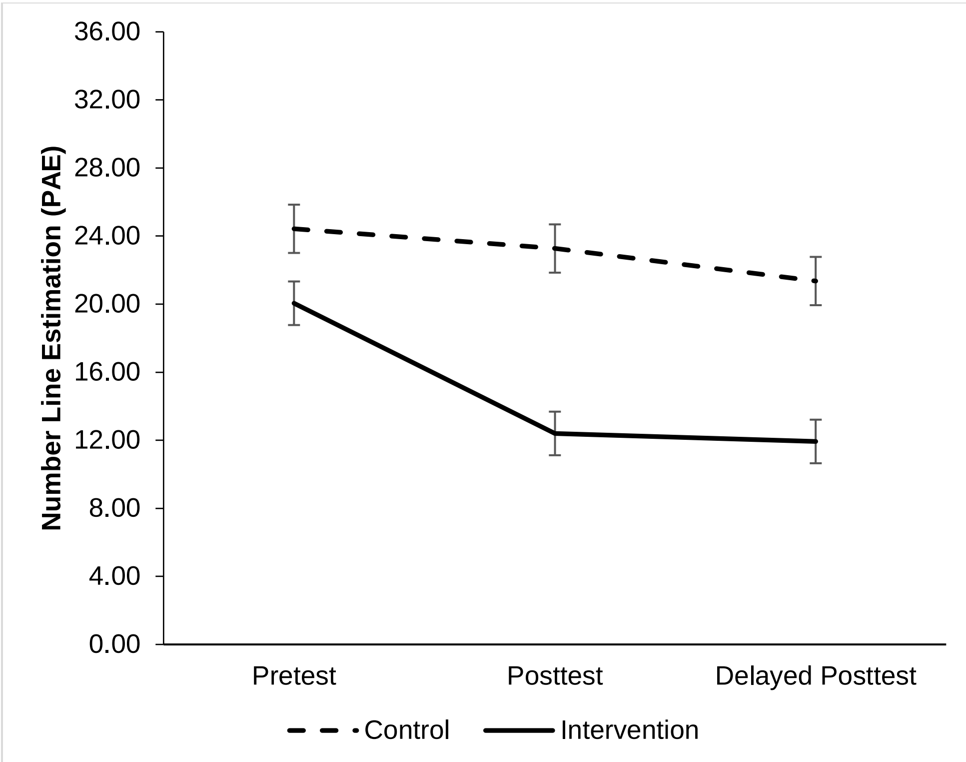

Figure 5. Estimated marginal means of Fraction number line estimation controlling for receptive vocabulary

Fraction Comparisons

As displayed in Table 7, there was no main effect of time, but a main effect of condition and a significant time by condition interaction. A trend analysis explicating the time by condition interaction revealed a trend toward a significant quadratic component but not a significant linear component. This trend is plotted in Figure 6, which displays an increase in scores from pretest to posttest for both conditions but a much greater increase for the experimental intervention group. The experimental intervention shows a minor decline in fraction comparison scores from post- to delayed posttest but still significantly outperforms the control condition. Post hoc pairwise comparisons revealed that the experimental intervention group had higher fraction comparison scores than the control condition at both immediate posttest (p = .001) and delayed posttest (p =

.005).

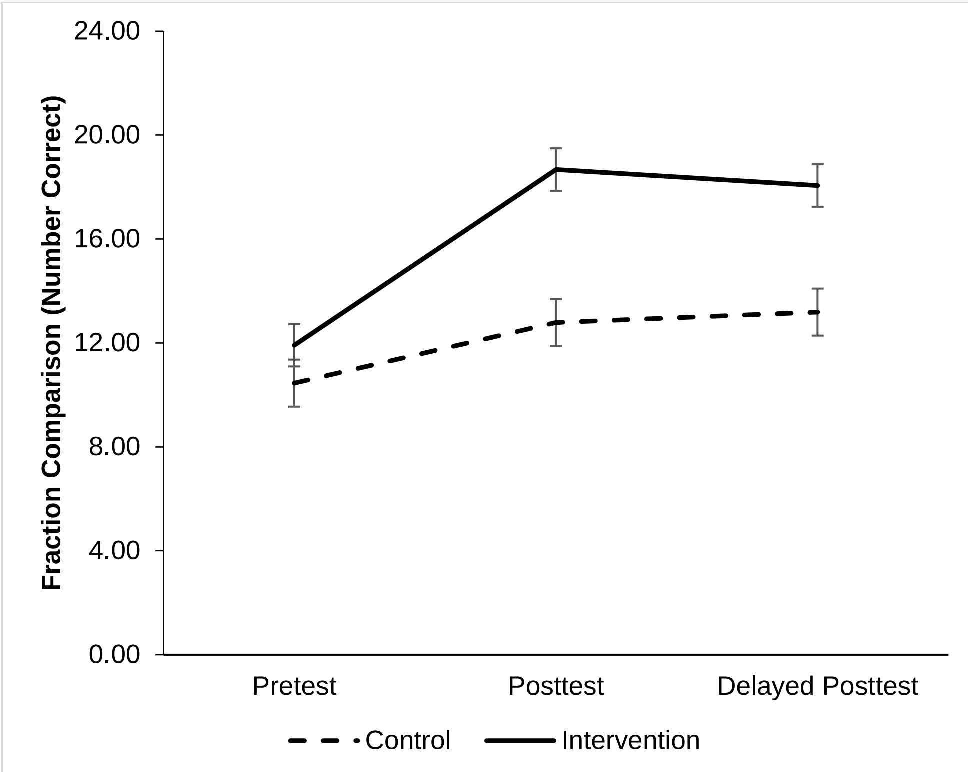

Figure 6. Estimated marginal means of fraction comparison controlling for receptive vocabulary

Fraction Arithmetic

As displayed in Table 8, there was no main effect of time or condition. There was no significant interaction between time and condition. Both conditions demonstrated comparably small improvements in fraction arithmetic scores between pretest and post- test and comparable declines between immediate and delayed posttest.

Effect Sizes

The experimental intervention yielded large effects on fraction concepts ($g = 1.09$), FNLE ($g = -0.85$), and fraction comparisons ($g = 0.82$) at posttest. Effects of the experimental intervention on these three measures at delayed posttest were medium to large (gs = 0.66, 0.60, 0.61, respectively). Effect sizes for fraction arithmetic were small at both posttests.

Exploration of Attentive Behavior as Moderator of the Effect of Intervention on Fraction Concepts

As displayed in Table 10 and in supplemental tables S11 and S12, the exploratory moderation models revealed a significant interaction between attentive behavior and intervention only when predicting posttest fraction concepts ($\beta = -.30$, $p = .048$). To explicate the significant interaction, simple effects were calculated. This was done with follow-up split regressions based on condition using the same variables included in the main effect model for fraction concepts. Posttest fraction concepts scores were regressed onto attentive behavior scores controlling for pretest fraction concepts. Results demonstrated that although attentive behavior scores were trending toward significance for the control condition ($\beta = .330$, $p = .085$), they were unrelated to posttest fraction concepts scores for the experimental intervention group ($\beta = -.020$, $p = .921$). These results suggest that on fraction concepts, students who normally demonstrated low attentive behavior in the classroom may have benefitted more from the experimental intervention than the regular school intervention provided to the business-as-usual control condition. This moderation is displayed in Figure 7, which displays the predicted fraction concepts posttest score of a student who was one standard deviation above the mean and a student who was one standard deviation below the mean on the SWAN scale.

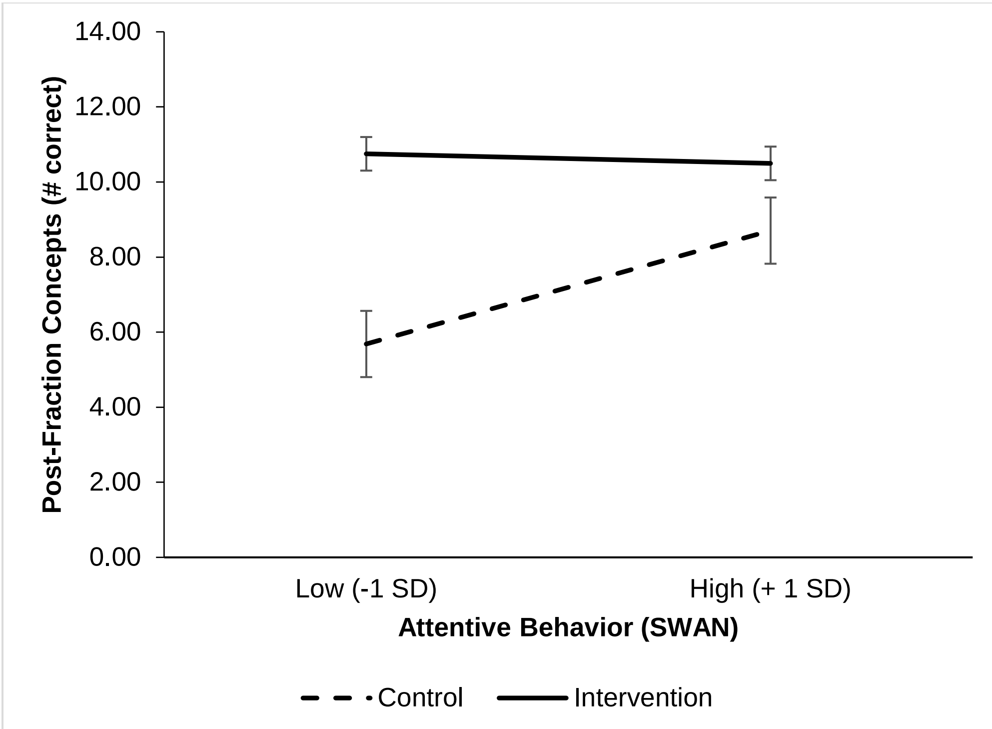

Figure 7. Moderating effect of attentive behavior on predicted fraction concepts score by condition

Discussion

We evaluated the effectiveness of an experimental intervention designed to improve fraction learning in sixth-grade students with mathematics difficulties. The experimental intervention builds on a growing body of research stressing the importance of number lines as representational tools for learning key mathematics concepts (Fazio et al., 2016; Siegler et al., 2010). Understanding that all real numbers can be represented as magnitudes on a number line provides a unifying structure for most mathematical learning, including fractions (Siegler et al., 2011). The experimental intervention was further supported by validated strategies from studies on the Science of Learning.

As predicted, the experimental intervention led to large and meaningful improvements on student measures aligned specifically to the experimental intervention (i.e., fraction number line estimation acuity and fraction magnitude comparisons) as well as on a broader measure of fraction concepts. Pre- to posttest effect sizes (Hedges' *g*) were large, ranging from 0.82 to 1.09. Importantly, these improvements generally held at a 7-week delay. Because of stratified random sampling and random assignment, we are confident that receipt of the experimental intervention was the only systematic difference between conditions. Conditions did not differ on fraction pretest measures, demographics, and cognitive measures (with the exception of vocabulary, which we controlled for in all subsequent analyses). Further, high instructor fidelity reflects consistent administration of the experimental intervention across instructors. Additionally, we used a validated screener (Rodrigues et al., 2017) to identify selected students at risk for mathematics failure, ensuring that our sample was highly targeted to increase the applicability of our current findings to the appropriate population.

Our findings add to the current literature on the effectiveness of a number-line approach to teaching fractions in important ways. First, the findings show that such an experimental intervention successfully boosts fraction knowledge in older students (i.e., sixth graders) whose mathematical difficulties are likely to be entrenched, relative to those of younger learners. Much of the previous fraction number line intervention work has focused on at-risk fourth graders (Fuchs et al., 2014) because this is the grade when formal fraction instruction first takes place. In the present study, we undertook a more challenging task. That is, we chose to focus on sixth-grade students who had already received several years of formal instruction on fractions, but who still showed weak performance on a validated fraction screener at sixth grade entry.

Mathematics achievement shows steep declines around the time that students transition into middle school (Wang & Pomerantz, 2009), and fractions are a key element of mathematics achievement during this period. According to the 2015 National Assessment of Educational

Progress (NAEP), only 40% of fourth graders in the U.S. are proficient in mathematics and this declines to 33% in eighth grade. Strikingly, without intervention, studies have shown that students with mathematics difficulties make minimal gains in fraction learning between sixth and eighth grades (Mazzocco & Devlin, 2008; Siegler & Pyke, 2013). As such, sixth grade is a critical time to catch students who may have fallen through the cracks in fraction learning during intermediate grades.

Second, experimental intervention students' gains were sustained after an almost two-month delay. We speculate that integration of strategies based on learning principles worked as intended, leading to deeper learning that was retained at a 7-week delay. Explicit instruction likely provided the appropriate teacher modeling and student practice necessary to build students' repertoire of problem-solving strategies (Doabler & Fien, 2013). Representational gestures encouraged students to form abstract representations of fractions, as suggested by Ping and Goldin-Meadow (2008), leading to improved performance on the three conceptually oriented fraction measures. Presenting information verbally and visually likely fostered dual coding and improved retention of the learning material (Cuevas, 2016) along with spaced and interleaved practice (Rohrer et al., 2015). Finally, it is probable that immediate and process-oriented feedback throughout the lessons necessary for learners with low prior knowledge encouraged students to refine their thinking and problem-solving skills (Graesser et al., 2005; Fyfe & Rittle-Johnson, 2016). However, it is important to note that we chose to employ these learning principles in combination, rather than isolating the effects of each feature. Our "engineering" approach aimed to maximize student learning through the use of established, research-based learning strategies. Future work would need to experimentally isolate these features to determine causal effects of each principle on the development of fraction understanding in particular.

Our sample demonstrated below-average performance on most of the cognitive and behavioral measures, including working memory, nonverbal reasoning, receptive vocabulary, and classroom attentive behavior. Even after employing random assignment strat ified by classroom, our intervention and control group showed significant differences on receptive vocabulary prior to the start of the intervention. When controlling for receptive vocabulary we found that vocabulary predicted performance on the general fraction concepts measure and number line task but did not differentially predict performance by condition. This is not surprising when considering the fraction concepts measure which includes several word problems. Vocabulary did not predict performance on the fraction comparison or arithmetic measures. Although not a focus of the current study, these findings may suggest a potentially interesting area for future work. A recent study demonstrated that syntactic ability was a significant predictor of first and second graders' mathematics performance whereas vocabulary was not (Chow & Ekholm, 2019). Yet, earlier work suggests that a range of oral language skills, including receptive vocabulary, are important predictors of fraction performance in particular (Chow & Jacobs, 2016). Future work should explore further the role that oral language skills, and vocabulary in particular, may play in learning fraction concepts.

One curious finding was that students demonstrated average inhibitory control as measured by the flanker task, which was not highly correlated with classroom attentive behavior (r = .289).[2] This low correlation suggests that these measures (i.e., flanker task and SWAN) assess related but distinct forms of attention. We propose that the measure of classroom attentive behavior used in the current study taps students' self-regulation skills, whereas the flanker task is an assessment of selective *visual* attention in particular. The exploratory analyses involving classroom attentive

[2] A correlation matrix between pretest fraction measures and cognitive competencies is available in a supplemental online appendix.

behavior suggests the SWAN measure as a potentially important measure of attention as described below.

A potentially interesting exploratory finding was the moderating effect of attentive behavior on condition. That is, students who exhibited inattentive behavior in their regular mathematics classroom (based on teacher reports) tended to benefit more from the experimental intervention than from the regular school intervention that was provided to the business-as-usual control group (i.e., working with the computer adaptive mathematics software). High attention in the classroom allows learners to stay on task and focus on key parts of a problem (Finn, Pannozzo, & Voelkl, 1995). Prior research has confirmed the direct and indirect influences of attentive behavior on broad fraction skills (e.g., Hecht et al., 2003). Classroom attention has been linked to specific fraction skills including fraction conceptual knowledge (Hansen et al., 2015), fraction magnitude estimation (Ye et al., 2016), and fraction arithmetic (Ye et al., 2016). Arguably, the explicit supports provided in our experimental intervention helped students attend more often than the school control intervention administered individually on computers. It is also possible that the small, synergistic groups that our experimental intervention provided were important in this respect. In any case, the results should be interpreted cautiously, considering that this moderating effect was present only on the broad fraction concepts measure—not on the number line estimation or comparison tasks. These exploratory findings should spark future systematic study with a larger sample on the relationship between classroom attentive behavior and the effects of fractions interventions.

Although our primary hypotheses received support, we did not find improvements in fraction arithmetic for the experimental intervention students. Rather, growth in fraction arithmetic followed a pattern similar to that of the control condition, with slight increases between pre- and posttest and then a slight decline at delayed posttest. It is possible that the minor changes in students' fraction arithmetic were associated with students' regular classroom instruction rather than differential experiences during the intervention period. The finding is not surprising given that our experimental intervention focused on developing fraction concepts, although some research suggests that instruction in fraction magnitudes leads to improved fraction arithmetic as well as concepts (Bailey, Hoard, Nugent, & Geary, 2012). However, a post hoc power analysis suggested we were underpowered for detecting small effects. Another issue is that our arithmetic assessment may not have been sufficiently sensitive to growth. Many of the items did not match the content of the limited arithmetic instruction provided in our experimental intervention, leading to few opportunities for students to demonstrate the arithmetic they may have learned. For example, although modeling multiplication of a fraction and a whole number was covered in two lessons of focused instruction and interleaved review over seven lessons, only three of the arithmetic items assessed this skill. Examination of performance at the item level showed that experimental intervention students answered three particular arithmetic items more accurately at posttest than did controls: two addition with common denominators and one multiplication (1/2 X 24). For each of these problems, 20% to 30% more of the students in the experimental intervention group answered correctly than those in the control condition. Future iterations of our experimental intervention should use measures that align fraction arithmetic problems more closely to the content that is taught as well as address the conceptual underpinnings of procedures taught in the classroom to help students avoid misapplication (Rittle-Johnson, 2017).

Our lack of specific information about the instruction offered to the business-as-usual control during their corresponding school intervention period as well as in regular mathematics classes limits interpretation of the results. Although we can confirm that students in the control condition were simultaneously receiving an individualized mathematics intervention offered by their schools, the software programs were adaptive. As such, the specific content each control student was exposed to likely varied, but this was not made available to us. We also were not given specific information on how fractions are normally taught in their regular mathematics classes, although both participating schools followed CCSS-M benchmarks (NGACBP & CCSSO, 2010) and used the same curriculum (*Connected Mathematics Project*). The sixth grade CCSS-M benchmarks emphasize understanding of rational numbers on a number line, but it is unclear to what extent our sample's regular mathematics teachers incorporated the number line into their regular instruction. The finding that children in both groups gained at the same rate in arithmetic but not on fraction magnitude concepts (e.g., fraction number line estimation) suggests that arithmetic may have been the focus in regular mathematics classes. Still, if differences did exist in the treatment of fractions in their classrooms, these differences would be randomly distributed across conditions and should not systematically influence the impact of our experimental intervention on students' fraction knowledge.

One interesting finding is that, although students in the experimental intervention condition outperformed the control group at delayed posttest on the broad fraction concepts measure (with medium to large effects), the control group showed a small but noticeable improvement between immediate and delayed posttest while the intervention condition levels off. Formal fraction instruction within students' regular mathematics class would have concluded at the point of immediate posttest. Even if fraction instruction continued in some of the regular mathematics classes, this would not have systematically differed by condition as students from both the intervention and control groups were distributed evenly within each regular mathematics class. One plausible explanation is that students in the control group as well as their parents, mathematics teachers, and intervention period teachers were each made aware of their need for mathematics intervention particularly in the area of fractions. Students were also aware that they would be taking the fraction assessment again. It is possible that the control group students' mathematics teachers assigned these students additional practice with fractions either during class or as homework between immediate and delayed posttest to make up for the lack of the comprehensive fraction intervention that the experimental students received. Unfortunately, we do not have access to detailed records of regular classroom instruction that occurred over the course of data collection for the current study. However, school administrators and teachers did not mention this issue to our research team. Future work should include collection of more detailed information on classroom instruction received over the course of the study to help interpret results more contextually.

The pattern of results for the broad fraction concepts measure suggests that the control group may eventually catch up to the intervention group. Future work should include one or more additional time points to assess longer-term retention of improvements by condition. Because of the lack of attention fractions receives in formal instruction within the regular sixth-grade mathematics class, students may need a longer-term intervention to show continued improvements, either in the form of a longer intervention period, additional refresher sessions scheduled later in the school year, or the same number of sessions but spread out across the school year. These possibilities should be explored further.

To understand the practical significance of the current findings, we examined how students in our experimental intervention group compared with their higher-performing peers at the completion of the intervention period. A previous longitudinal study with a sample representative of the population that the current study participants were sampled from found three distinct growth trajectories in fraction number line estimation skills (Resnick et al., 2016), using a similar FNLE measure. As noted earlier, one group of students in the 2016 study started accurate in the fourth grade and ended accurate in the sixth grade (Class 1). These students were more likely to have high sixth grade mathematics achievement scores. Two groups of students started off with inaccurate number line estimates but one group showed steep improvements (Class 2) while the other showed minimal growth (Class 3). The group that started off inaccurate but showed substantial gains between fourth and sixth grades was more likely to have average sixth-grade mathematics achievement scores. The group that started off inaccurate and showed minimal growth was more likely to have low sixth grade mathematics achievement scores. In the present study sample, students made inaccurate number line estimates prior to the start of the intervention, with scores similar to the two inaccurate classes in the aforementioned 2016 longitudinal study. However, at the completion of the present intervention and at delayed posttest, the experimental intervention group showed steep growth with scores similar to Class 2, which was the group that was more likely to have average mathematics achievement scores at the end of sixth grade. The control condition showed no growth and had scores similar to Class 1 in the aforementioned 2016 study, which was the group that was likely to have low mathematics achievement scores. We suspect that our experimental intervention helps minimize the mathematics achievement gap between low and average-performing students, although further work is needed to confirm this suggestion.

Overall, our intervention findings hold promise for strengthening fraction skills in struggling students, which then should bolster later mathematics achievement more generally (Bailey et al., 2012) and algebra readiness in particular (Booth & Newton, 2012). We saw large improvements in fraction concepts, fraction number line estimation, and fraction comparison skills after just six weeks of targeted instruction, and these improvements held seven weeks later. Translating research on learning principles into real-world classrooms for struggling students while keeping the number line representation of fractions at the forefront seemed to have a powerful effect on the development of foundational mathematics knowledge. Perhaps the most important lesson to be learned from this study is that at-risk students can show marked improvements in their fraction knowledge several years after receiving initial instruction, if provided with a purposeful intervention designed to capitalize on learning principles from cognitive science and mathematics development.

References

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic & Physiological Optics, 34,* 502–508. http://dx.doi.org/10.1111/opo
.12131

Bailey, D. H., Fuchs, L. S., Gilbert, J. K., Geary, D. C., & Fuchs, D. (2018). Prevention: Necessary but insufficient? A 2-year follow-up of an effective first-grade mathematics intervention. *Child Development*.

Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology, 113,* 447– 455. http://dx.doi.org/10
.1016/j.jecp.2012.06.004

Bailey, D. H., Nguyen, T., Jenkins, J. M., Domina, T., Clements, D. H., & Sarama, J. S. (2016). Fadeout in an early mathematics intervention: Constraining content or preexisting differences? *Developmental Psychology, 52,* 1457–1469. http://dx.doi.org/10.1037/dev0000188

Bailey, D. H., Siegler, R. S., & Geary, D. C. (2014). Early predictors of middle school fraction knowledge. *Developmental Science, 17*, 775– 785. http://dx.doi.org/10.1111/desc.12155

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

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128,* 612– 637. http://dx.doi.org/10.1037/0033-2909.128.4.612

Bauch, E. M., & Otten, L. J. (2012). Study-test congruency affects encoding- related brain activity for some but not all stimulus materials. *Journal of Cognitive Neuroscience, 24,* 183–195. http://dx.doi.org/10.1162/ jocn_a_00070

Bauer, P. J., & Zelazo, P. D. (2013). IX. NIH Toolbox Cognition Battery (CB): Summary, conclusions, and implications for cognitive development. *Monographs of the Society for Research in Child Development, 78,* 133–146. http://dx.doi.org/10.1111/mono.12039

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B.* Methodological, *57,* 289 –300. http://dx.doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology, 46,* 545–551. http://dx.doi.org/10.1037/a0017887

Bonferroni, C. E. (1936). Statistical theory of classification and calculation of the probability. *Publication of the R Institute Superiore of the Science of Economy and Commerce of Florence, 8,* 3– 62.

Booth, J. L., McGinn, K. M., Barbieri, C., Begolli, K., Chang, B., Miller- Cotto, D., . . . Davenport, J. L. (2017). Evidence for cognitive science principles that impact learning in mathematics. In D. C. Geary & D. Berch (Eds.), *Mathematical cognition and learning* (Vol. 3, pp. 297– 325). London, UK: Academic Press. http://dx.doi.org/10.1016/B978-0- 12-805086-6.00013-8

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

Booth, J. L., 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. http://dx.doi

.org/10.1016/j.jecp.2013.09.001

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42,* 189 –201. http://dx.doi.org/10.1037/0012-1649.41.6.189

Bottge, B. A., Ma, X., Gassaway, L., Toland, M. D., Butler, M., & Cho,

S. J. (2014). Effects of blended instructional models on math performance. *Exceptional Children, 80,* 423– 437. http://dx.doi.org/10.1177/ 0014402914527240

Boyer, T. W., & Levine, S. C. (2012). Child proportional scaling: Is 1/3 = 2/6 = 3/9 = 4/12? *Journal of Experimental Child Psychology, 111,* 516 –533. http://dx.doi.org/10.1016/j.jecp.2011.11.001

Brown, P. C., Roediger, H. L., III, & McDaniel, M. A. (2014). *Make it stick: The science of successful learning*. Cambridge, MA: Belknap Press of Harvard University Press.

http://dx.doi.org/10.4159/978067 4419377

Bryant, D. P., Bryant, B. R., Gersten, R. M., Scammacca, N. N., Funk, C., Winter, A., . . . Pool, C. (2008). The effects of tier 2 intervention on the mathematics performance of first-grade students who are at risk for mathematics difficulties. *Learning Disability Quarterly, 31,* 47–63. http://dx.doi.org/10.2307/20528817

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry, 46,* 3–18. http://dx.doi.org/10.1111/ j.1469-7610.2004.00374.x

Butterworth, B., & Reigosa-Crespo, V. (2007). Information processing deficits in dyscalculia. In D. B. Berch & M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins or mathematical learning difficulties and disabilities* (pp. 65– 81). Baltimore, MD: Brookes.

Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review, 24,* 369 –378. http://dx.doi.org/10.1007/s10648-012-9205-z

Chen, X. (2009). *Students who study science, technology, engineering, and mathematics (STEM) in postsecondary education. Stats in brief*. NCES 2009 –161. Washington, DC: National Center for Education Statistics.

Chow, J. C., & Ekholm, E. (2019). Language domains differentially predict mathematics performance in young children. *Early Childhood Research Quarterly, 46,* 179 –186. http://dx.doi.org/10.1016/j.ecresq.2018.02.011 Chow, J. C., & Jacobs, M. (2016). The role of language in fraction performance: A synthesis of literature. *Learning and Individual Differ-ences, 47,* 252–257. http://dx.doi.org/10.1016/j.lindif.2015.12.017 Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of

gesture in bilingual education: Does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism, 7,* 303–319. http://dx.doi.org/10.1080/13670050408667815

Clarke, P., Crawford, C., Steele, F., & Vignoles, A. (2010). *The choice between fixed and random effects models: Some considerations for educational research. IZA Discussion Paper Series 5287*. Bonn, Germany: Institute for the Study of Labor.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cook, S. W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development, 7,* 211–232. http://dx.doi.org/10.1207/s15327647 jcd0702_4

Cuevas, J. (2016). An analysis of current evidence supporting two alternate learning models: Learning styles and dual coding. *Journal of Educational Sciences and Psychology, 6,* 1–13.

Cummings, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.

Cummings, K. D., & Smolkowski, K. (2015). Selecting students at risk of academic difficulties. *Assessment for Effective Intervention, 41,* 55– 61. http://dx.doi.org/10.1177/1534508415590396

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44,* 1– 42. http://dx.doi.org/10.1016/0010-0277(92)90049-N

DeWolf, M., & Vosniadou, S. (2015). The representation of fraction magnitudes and the whole number bias reconsidered. *Learning and Instruction, 37,* 39 – 49. http://dx.doi.org/10.1016/j.learninstruc.2014.07.002

Doabler, C. T., & Fien, H. (2013). Explicit mathematics instruction: What teachers can do for teaching students with mathematics difficulties. *Intervention in School and Clinic, 48,* 276 –285. http://dx.doi.org/10.1177/1053451212473151

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84,* 795– 805. http://dx.doi.org/10.1037/0021-9010.84.5.795

Dreambox Learning [Computer software]. (2012). Bellevue, WA: Dreambox Learning, Inc.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham,

D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14,* 4 –58. http://dx.doi.org/10.1177/1529100612453266

Dyson, N., Jordan, N. C., Rodrigues, J., Barbieri, C. A., & Rinne, L. (2018). A fraction sense intervention for students with or at risk for mathematics difficulties. *Remedial and Special Education*. Advance online publication. http://dx.doi.org/10.1177/0741932518807139

Empson, S. B., & Levi, L. (2011). *Extending children's mathematics: Fractions and decimals*. Portsmouth, NH: Heinemann.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http:// dx.doi.org/10.3758/BF03193146

Fazio, L. K., Kennedy, C. A., & Siegler, R. S. (2016). Improving children's knowledge of fraction bagnitudes. *PLoS ONE, 11,* 1–14. http://dx.doi

.org/10.1371/journal.pone.0165243

Finn, J. D., Pannozzo, G. M., & Voelkl, K. (1995). Disruptive and inattentive-withdrawn behavior and achievement among fourth graders. *The Elementary School Journal, 95,* 421–434. http://dx.doi.org/10.1086/461853

Fuchs, L., Compton, D., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett,

C. L. (2005). The prevention, identification, and cognitive determinants of mathematics difficulty. *Journal of Educational Psychology, 97,* 493– 513. http://dx.doi.org/10.1037/0022-0663.97.3.493

Fuchs, L. S., Malone, A. S., Schumacher, R. F., Namkung, J., & Wang, A. (2017). Fraction intervention for students with mathematics difficulties: Lessons learned from five randomized controlled trials. *Journal of Learning Disabilities*, *50,* 631– 639. http://dx.doi.org/10.1177/0022219416677249

Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., . . . Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology, 105,* 683–700. http://dx.doi.org/10.1037/a0032446

Fuchs, L. S., Schumacher, R. F., Malone, A. S., & Fuchs, D. (2011). *Fraction face-off*. Available from LS Fuchs, Vanderbilt University, Nashville, TN.

Fuchs, L. S., Schumacher, R. F., Sterba, S. K., Long, J., Namkung, J., Malone, A.,… Changas, P. (2014). Does working memory moderate the effects of fraction intervention? An aptitude–treatment interaction. *Journal of Educational Psychology*, *106,* 499 –514. http://dx.doi.org/10.1037/a0034341

Fyfe, E. R., & Rittle-Johnson, B. (2016). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology, 108,* 82–97. http://dx.doi.org/10.1037/edu0000053

Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology, 104*(4), 1094.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37,* 4 –15. http://dx.doi.org/10.1177/00222194040370010201

Gershon, R. C., Wagster, M. V., Hendrie, H. C., Fox, N. A., Cook, K. F., & Nowinski, C. J. (2013). NIH toolbox for assessment of neurological and behavioral function. *Neurology, 80,* S2–S6. http://dx.doi.org/10.1212/WNL.0b013e3182872e5f

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools. NCEE 2009 – 4060. *What Works Clearinghouse*. Retrieved from http:// files.eric.ed.gov/fulltext/ED504995.pdf

Gersten, R., Schumacher, R. F., & Jordan, N. C. (2017). Life on the number line: Routes to understanding fraction magnitude for students with difficulties learning mathematics. *Journal of Learning Disabilities, 50,* 655– 657. http://dx.doi.org/10.1177/0022219416662625

Goldin-Meadow, S. (2011). Learning through gesture. *WIREs Cognitive Science, 2,* 595– 607. http://dx.doi.org/10.1002/wcs.132

Graesser, A. C., McNamara, D. S., & VanLehn, K. (2005). Scaffolding deep comprehension strategies through Point& Query, AutoTutor, and iSTART. *Educational Psychologist, 40,* 225–234. http://dx.doi.org/10.1207/s15326985ep4004_4

Hamdan, N., & Gunderson, E. A. (2017). The number line is a critical spatial-numerical representation: Evidence from a fraction intervention. *Developmental Psychology, 53*(3), 587.

Handel, M. J. (2016). What do people do at work? A profile of U.S. jobs from the survey of workplace Skills, Technology, and Management Practices (STAMP). *Journal of Labour Market Research, 49,* 177–197. http://dx.doi.org/10.1007/s12651-016-0213-1

Hanover Research. (August 2014). *Best practices in math interventions*. Retrieved from https://www.mbaea.org/media/cms/Best_Practices_in_Math_Intervention_53D80FEED7650.pdf

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. http://dx.doi.org/10.1016/j.cogdev.2015.02.001

Hecht, S. A., Close, L., & Santisi, M. (2003). Sources of individual differences in fraction skills. *Journal of Experimental Child Psychology, 86,* 277–302. http://dx.doi.org/10.1016/j.jecp.2003.08.003

Huang, F. L. (2016). Alternatives to multilevel modeling for the analysis of clustered data. *Journal of Experimental Education, 84,* 175–196. http://dx.doi.org/10.1080/00220973.2014.952397

i-Ready [Computer software]. (2016). North Billerica, MA: Curriculum Associates.

Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). Arithmetic fact mastery in young children: A longitudinal investigation. *Journal of Experimental Child Psychology, 85,* 103–119. http://dx.doi.org/10.1016/S0022-0965(03)00032-8

Jordan, N. C., Resnick, I., Rodrigues, J., Hansen, N., & Dyson, N. (2017). Delaware longitudinal study of fraction learning: Implications for helping children with mathematics difficulties. *Journal of Learning Disabilities, 50,* 621– 630. http://dx.doi.org/10.1177/0022219416662033

Kennedy, P. E. 2003. *A guide to econometrics*. Cambridge, MA: MIT Press. IBM Corp. (2016). IBM SPSS statistics for Windows (Version 24.0). Armonk, NY: Author.

Koedinger, K. R., Booth, J. L., & Klahr, D. (2013). Education research. Instructional complexity and the science to constrain it. *Science*, *342*, 935–937. http://dx.doi.org/10.1126/science.1238056

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4,* 863. http://dx.doi.org/10.3389/fpsyg.2013.00863

Lakes, K. D., Swanson, J. M., & Riggs, M. (2012). The reliability and validity of the English and Spanish Strengths and Weaknesses of ADHD and Normal behavior rating scales in a preschool sample: Continuum measures of hyperactivity and inattention. *Journal of Attention Disorders, 16,* 510 –516. http://dx.doi.org/10.1177/1087054711413550 Lappan, G., Difanis Phillips, E., Fey, J., & Friel, S. (2014). *Connected*

Mathematics 3, Grade 6. Upper Saddle River, NJ: Pearson Education. Malone, A. S., & Fuchs, L. S. (2017). Error patterns in ordering fractions among at-risk fourth-grade students. *Journal of Learning Disabilities, 50,* 337–352. http://dx.doi.org/10.1177/0022219416629647

Mason, B. J., & Bruning, R. (2001). Providing feedback in computer-based instruction: What the research tells us. Center for Instructional Innovation, University of Nebraska-Lincoln: 14. (Tech. Rep.).

Matthews, M. S., & Farmer, J. L. (2008). Factors affecting the algebra I achievement of academically talented learners. *Journal of Advanced Academics, 19,* 472–501. http://dx.doi.org/10.4219/jaa-2008-810

Mayfield, K. H., & Chase, P. N. (2002). The effects of cumulative practice on mathematics problem solving. *Journal of Applied Behavior Analysis, 35,* 105–123. http://dx.doi.org/10.1901/jaba.2002.35-105

Mazzocco, M. M. M., & Devlin, K. T. (2008). Parts and 'holes': Gaps in rational number sense among children with vs. without mathematical learning disabilities. *Developmental Science, 11,* 681– 691. http://dx.doi.org/10.1111/j.1467-7687.2008.00717.x

Mix, K. S., Levine, S. C., & Huttenlocher, J. (1999). Early fraction calculation ability. *Developmental Psychology, 35,* 164 –174. http://dx.doi.org/10.1037/0012-1649.35.1.164

Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science, 32,* 99 –113. http://dx.doi.org/10.1023/B: TRUC.0000021811.66966.1d

Murrar, S., & Brauer, M. (2018). Mixed model analysis of variance. In B. B. Frey (Ed.), *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation* (pp. 1075–1078). Thousand Oaks, CA: SAGE Publications. http://dx.doi.org/10.4135/9781506326139.n436

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core state standards for mathematics*. Washington, DC: Author.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U. S. Department of Education.

Opfer, J. E., & Devries, J. M. (2008). Representational change and magnitude estimation: Why young children can make more accurate salary comparisons than adults. *Cognition, 108,* 843– 849. http://dx.doi.org/10.1016/j.cognition.2008.05.003

Paivio, A. (1986). *Mental representations: A dual-coding approach*. Oxford, UK: Oxford University Press.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal, 316,* 1236 –1238. http://dx.doi.org/10.1136/bmj.316.7139.1236

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, *3*, 359 – 400. http://dx.doi.org/10.1016/0885-2014(88)90021-4

Ping, R. M., & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology, 44,* 1277–1287. http://dx.doi.org/10.1037/0012-1649.44.5.1277

Rea, C. P., & Modigliani, V. (1985). The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. Human Learning. *Human Learning: Journal of Practical Research & Applications, 4,* 11–18.

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. http://dx.doi.org/10.1037/dev0000102

Rittle-Johnson, B. (2017). Developing mathematics knowledge. *Child Development Perspectives, 11,* 184 –190. http://dx.doi.org/10.1111/cdep.12229

Rittle-Johnson, B., & Jordan, N. C. (2016). *Synthesis of IES-Funded Research on Mathematics: 2002–2013*. NCER 2016 –2003. Washington, DC: National Center for Education Research.

Rodrigues, J., Dyson, N. I., Hansen, N., & Jordan, N. C. (2016). Preparing for algebra by building fraction sense. *Teaching Exceptional Children, 49,* 134 –141. http://dx.doi.org/10.1177/0040059916674326

Rodrigues, J., Jordan, N. C., Hansen, N., Resnick, I., & Ye, A. (2017, October). Identifying fraction concepts measures as effective screeners of mathematics risk status. Poster presented at the meeting of the Cognitive Development Society (CDS), Portland, OR.

Rohrer, D., Dedrick, R. F., & Stershic, S. (2015). Interleaved practice improves mathematics learning. *Journal of Educational Psychology*, *107*, 900 –908. http://dx.doi.org/10.1037/edu0000001

Rohrer, D., & Taylor, K. (2006). The effects of overlearning and distributed practice on the retention of mathematics knowledge. *Applied Cognitive Psychology, 20,* 1209 –1224. http://dx.doi.org/10.1002/acp.1266

Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science, 35,* 481– 498. http://dx.doi.org/ 10.1007/s11251-007-9015-8

Saxe, G. B., Diakow, R., & Gearhart, M. (2013). Towards curricular coherence in integers and fractions: A study of the efficacy of a lesson sequence that uses the number line as the principal representational context. *ZDM, 45*, 343–364. http://dx.doi.org/10.1007/s11858-012-0466-2

Schneider, B., Swanson, C. B., & Riegle-Crumb, C. (1998). Opportunities for learning: Course sequences and positional advantages. *Social Psychology of Education, 2,* 25–53. http://dx.doi.org/10.1023/A:1009601517753

Schneider, M., Thompson, C. A., & Rittle-Johnson, B. (2018). Associations of magnitude comparison and number line estimation with mathematical competence: A comparative review. In P. Lemaire (Ed.), *Cognitive development from a strategy perspective: A festschrift for Rober S.* Siegler (pp. 100 –119). London, UK: Routledge.

Shin, M., & Bryant, D. P. (2017). Improving the fraction word problem solving of students with mathematics learning disabilities: Interactive computer application. *Remedial and Special Education, 38,* 76 – 86. http://dx.doi.org/10.1177/0741932516669052

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78,* 153–189.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children.

Child Development, 75, 428 – 444. http://dx.doi.org/ 10.1111/j.1467-8624.2004.00684.x

Siegler, R. S., Carpenter, T., Fennell, F., Geary, D., Lewis, J., Okamoto, Y., . . . Wray, J. (2010). *Developing effective fractions instruction for kindergarten through 8th grade: A practice guide (NCEE #2010 – 4039)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U. S. Department of Education.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14,* 237–250. http://dx.doi.org/10.1111/1467-9280.02438

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

Siegler, R. S., Thompson, C. A., & Opfer, J. E. (2009). The logarithmic-to-linear shift: One learning sequence, many tasks, many time scales. *Mind, Brain, and Education, 3,* 143–150. http://dx.doi.org/10.1111/j.1751-228X.2009.01064.x

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

Swanson, J., Schuck, S., Mann, M., Carlson, C., Hartman, K., Sergeant, J.,. . . Wigal, T. (2006). *Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and SWAN rating scales*. Irvine, CA: University of California.

Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology, 24,* 837– 848. http://dx.doi.org/10.1002/ acp.1598

Thompson, C. A., & Opfer, J. E. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational changes in numerical cognition. *Child*

Development, 81, 1768 –1786. http://dx.doi.org/10.1111/j.1467-8624.2010.01509.x

U.S. Department of Education, Institute of Education Sciences. (2014). What works clearinghouse: Procedures and standards handbook (Version 3. 0). Retrieved from https://ies.ed.gov/ncee/wwc/Docs/reference resources/wwc_procedures_v3_0_standards_handbook.pdf

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

Wang, Q., & Pomerantz, E. M. (2009). The motivational landscape of early adolescence in the United States and China: A longitudinal investigation. *Child Development, 80,* 1272–1287. http://dx.doi.org/10.1111/j.1467-8624.2009.01331.x

Wechsler, D. (1999). *Manual for the Wechsler abbreviated intelligence scale (WASI).* San Antonio, TX: The Psychological Corporation.

Welcome, S. E., Paivio, A., McRae, K., & Joanisse, M. F. (2011). An electrophysiological study of task demands on concreteness effects: Evidence for dual coding theory. *Experimental Brain Research, 212,* 347–358. http://dx.doi.org/10.1007/s00221-011-2734-8

Ye, A., Resnick, I., Hansen, N., Rodrigues, J., Rinne, L., & Jordan, N. C. (2016). Pathways to fraction learning: Numerical abilities mediate the relation between early cognitive competencies and later fraction knowledge. *Journal of Experimental Child Psychology, 152,* 242–263. http:// dx.doi.org/10.1016/j.jecp.2016.08.001

Received June 1, 2018

Revision received March 25, 2019

Accepted May 13, 2019 ■

Table 1. *Scope and sequence of the fraction intervention*

Key Topics	Lessons 1-4 Halves	Lessons 5-9 Halves Fourths	Lessons 10-15 Halves Fourths Eighths	Lessons 20-21 Thirds Sixths Twelfths	Lessons 22-24 All studied denominators
Counting by unit fractions, by whole and mixed numbers	●	●	●		
Partitioning using linear, area, and set models	●	●	●	●	
Finding 1/b of a set (Multiplication)	●	●	●		
Finding a/b of a set when a ≠1 (Multiplication)	●	●	●		
Adding/Subtracting fractions with common denominators	●	●	●		
Locating mixed numbers on the number line	●	●	●		
Equating to improper fractions and mixed numbers	●	●	●		
Measuring with rulers marked with whole numbers and…	●	●	●		
Finding equivalent fractions with different denominators		●	●		
Comparing fraction magnitudes using various strategies[a]			●	●	
Measuring with cups[b]				●	
Adding/Subtracting with unlike denominators					●

[a] Activities involving comparison of fraction magnitudes provided opportunities for students to apply fraction comparison strategies (e.g., benchmark strategy) to denominators not studied within the intervention, such as fifths.
[b] Involved only denominators common to measuring cups: halves, thirds, fourths, and eighths.

Table 2. *Overview of lesson structure*

Activity	Description	Time
Warm-up	Individual worksheet practice of material from previous day.	3 minutes
Multiplication practice	Speeded practice of whole number multiplication facts using multiplicands that are aligned with denominators in the corresponding lesson.	3 minutes
Counting	Practice of oral counting of fractions with like denominators (e.g., "one-fourth, two-fourths, three-fourths…") using the number line as reference.	3 minutes
Targeted instruction	Explicit instruction targeting the lesson's learning goals and focused on the number line.	20 - 25min.
Games	Short, fast-paced card games targeting fraction magnitude judgements (e.g., comparing two fractions to each other, to one-half) and fraction equivalencies (e.g., 3 is the same as how many halves?; 3 halves is the same as how many fourths?).	3 minutes
Cool Down (Independent Practice and Formative Assessment)	Individual worksheet practice of material from that day's lesson and prior content.	3 minutes

Table 3. *Descriptives and tests of equivalence between conditions at pretest*

	Business-as-Usual Control (n = 23)		Intervention Group (n = 28)		Difference test
Continuous variables	Mean	*SD*	Mean	*SD*	sig. *T*-test
Fraction concepts	9.83	3.38	9.54	1.86	$t = 0.369$, $p = .715$, *ns*
Fraction arithmetic	3.65	2.62	4.07	3.03	$t = -0.522$, $p = .604$, *ns*
Fraction comparisons	10.35	3.80	12.00	5.73	$t = -1.185$, $p = .242$, *ns*
FNLE (PAE)	24.25	9.71	20.20	7.32	$t = 1.70$, $p = .096$, *ns*
Attentive behavior	35.65	13.39	34.04	11.02	$t = 0.473$, $p = .638$, *ns*
Receptive vocabulary	81.37	6.99	77.55	5.60	$t = 2.166$, $p = .035$*
Working memory	86.86	11.50	90.08	10.93	$t = -1.009$, $p = .318$, *ns*
Inhibitory control	102.53	14.08	103.05	13.41	$t = -0.133$, $p = .894$, *ns*
Nonverbal reasoning	33.83	10.19	36.43	9.44	$t = -0.945$, $p = .349$, *ns*
Proportional reasoning	14.36	5.34	16.11	5.56	$t = -1.120$, $p = .268$, *ns*
Categorical variables	%		%		sig. χ^2
Male	34.78%		50%		$\chi^2 = 1.192$, $p = .275$, *ns*
Special Education (Spec Ed)	30.43%		32.14%		$\chi^2 = 0.017$, $p = .896$, *ns*
English Language Learner (ELL)	13.04%		7.14%		$\chi^2 = 0.497$, $p = .481$, *ns*
Math Learning Disability (MLD)	30.43%		25.00%		$\chi^2 = 0.187$, $p = .665$, *ns*
Reading Learning Disability (RLD)	30.43%		28.57%		$\chi^2 = 0.021$, $p = .884$, *ns*
Behavioral Disability	0.00%		3.57%		$\chi^2 = 0.838$, $p = .360$, *ns*

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

Note: Fraction concepts, fraction arithmetic, and fraction comparison scores are sum scores with the highest possible score of 24. FNLE is percent absolute error with lower scores indicating higher accuracy. Attentive behavior is a sum score ranging from 9 – 63. Receptive vocabulary, working memory, and inhibitory control are scaled scores (M = 100, SD = 15). Nonverbal reasoning is a T-score (M = 50, SD = 10). Proportional reasoning is number correct of 24 trials.

Table 4. *Estimated marginal means (EMM) for fraction outcome measures controlling for receptive vocabulary*

	Pretest				Posttest				Delayed Posttest			
	Control		Intervention		Control		Intervention		Control		Intervention	
	EMM	*SE*	*EMM*	*SE*	*EMM*	*SE*	*EMM*	*SE*	*EMM*	*SE*	*EMM*	*SE*
Fraction Concepts	9.45	0.52	9.85	0.47	11.09	0.65	15.14	0.59	12.20	0.70	14.84	0.63
FNLE	24.43	1.83	20.05	1.65	23.27	1.55	12.40	1.40	21.36	1.69	11.93	1.53
Fraction Comparison	10.45	1.07	11.91	0.96	12.79	1.16	18.68	1.05	13.19	1.21	18.06	1.09
Fraction Arithmetic	3.46	0.60	4.23	0.55	6.94	1.22	8.66	1.10	5.84	0.89	7.06	0.81

Note: Adjusted for vocabulary at M = 79.27

All scores are raw scores with the exception of fraction number line estimation (FNLE), which uses the percent absolute error (PAE) where lower scores indicate better performance.

Table 5. *Mixed 2 (Condition) x 3 (Time) ANCOVA on pre-, post-, and delayed posttest Fraction Concepts scores by condition (intervention vs. control) controlling for vocabulary*

Source	*df*	*F*	*p*	*MSE*	η_p^2
Condition	1, 48	10.53	.002	18.32	.18
Time	1.74, 83.71	0.49	.611	4.28	.01
Vocabulary	1, 48	11.85	.001	18.32	.20
Time × Vocabulary	1.74, 83.71	0.37	.662	4.28	.01
Time × Condition	1.74, 83.71	10.41	<.001	4.28	.18
Within-subject contrasts: Time x Condition interaction					
Linear	1, 48	6.31	.015	4.59	.12
Quadratic	1, 48	16.93	<.001	2.89	.26

Note. MSE = mean square error

Table 6. *Mixed 2 (Condition) x 3 (Time) ANCOVA on pre-, post-, and delayed posttest Fraction Number Line Estimation (PAE) by condition (intervention vs. control) controlling for vocabulary*

Source	*df*	*F*	*p*	*MSE*	η_p^2
Condition	1, 48	17.70	<.001	132.06	.27
Time	1.72, 82.70	2.17	.128	32.39	.04
Vocabulary	1, 48	4.70	.035	132.06	.09
Time × Vocabulary	1.72, 82.70	3.20	.053	32.39	.06
Time × Condition	1.72, 82.70	4.80	.014	32.39	.09
Within-subject contrasts: Time x Condition interaction					
Linear	1, 48	3.81	.057	38.65	.07
Quadratic	1, 48	7.05	.011	17.15	.13

Note. MSE = mean square error

Table 7. *Mixed 2 (Condition) x 3 (Time) ANCOVA on pre-, post-, and delayed posttest Fraction Comparisons by condition (intervention vs. control) controlling for vocabulary*

Source	*df*	*F*	*p*	*MSE*	η_p^2
Condition	1, 48	10.69	.002	53.65	.18
Time	2, 96	1.21	.301	16.47	.03
Vocabulary	1, 48	0.04	.837	53.65	<.01
Time × Vocabulary	2, 96	1.29	.282	16.47	.03
Time × Condition	2, 96	3.77	.027	16.47	.07
Within-subject contrasts: Time x Condition interaction					
Linear	1, 48	3.59	.064	18.73	.07
Quadratic	1, 48	4.01	.051	14.70	.08

Note. MSE = mean square error

Table 8. *Mixed 2 (Condition) x 3 (Time) ANCOVA on pre-, post-, and delayed posttest Fraction Arithmetic by condition (intervention vs. control) controlling for vocabulary*

Source	*df*	*F*	*p*	*MSE*	η_p^2
Condition	1, 48	1.25	.269	42.11	.03
Time	2, 96	0.02	.982	7.85	< .01
Vocabulary	1, 48	2.22	.143	42.11	.04
Time × Vocabulary	2, 96	0.26	.769	7.85	.01
Time × Condition	2, 96	0.33	.719	7.85	.01
Within-subject contrasts: Time x Condition interaction					
Linear	1, 48	.16	.689	7.54	<.01
Quadratic	1, 48	.49	.489	8.16	.01

Note. MSE = mean square error

Table 9. *Effect sizes on differences between estimated marginal means between conditions*

	Pre - Posttest			Pre - Delayed Posttest		
	g	U3	Improvement Index	*g*	U3	Improvement Index
Fraction Concepts	1.09	.87	36.74%	0.66	.75	24.89%
Fraction NLE (PAE)	-0.85	.19	30.81%	-0.60	.27	23.01%
Fraction Comparison	0.82	.80	29.76%	0.61	.73	23.17%
Fraction Arithmetic	0.17	.57	6.71%	0.11	.54	4.42%

Note. Comparisons of estimated marginal means controlling for receptive vocabulary. Fraction number line estimation (FNLE) is scored as percent absolute error (PAE) where lower scores indicate better performance.

Table 10. *Prediction models of posttest fraction concepts scores*

	Fraction concepts							
	Pretest model		Main effects model		Covariate model		Moderation model	
	β	*B* *(SE)*	β	*B* *(SE)*	β	*B* *(SE)*	β	*B* *(SE)*
Pretest fraction concepts	.47**	0.65 (0.18)	.49**	0.67 (0.15)	.42**	0.58 (0.16)	.35**	0.48 (0.17)
Intervention			.46**	3.38 (0.80)	.47**	3.44 (0.79)	.47**	3.43 (0.76)
Attentive behavior					.18	0.05 (0.04)	.41*	0.13 (0.05)
Attentive behavior × Intervention							-.30*	-0.14 (0.07)
R^2	.22		.43		.46		.50	
ΔR^2 from prior model			.21**		.03		.05*	
F	13.54**		18.16**		13.22**		11.61**	

$p < .05$ level. * $p < .01$ level. **