# Supported Self-Explaining during Fraction Intervention

Lynn S. Fuchs[1], Amelia Malone[1], Robin F. Schumacher[1],

Jessica Namkung[2], Carol L. Hamlett[1],

Nancy C. Jordan[3], Robert S. Siegler[4], Russell Gersten[5], and Paul Changas[6]

Vanderbilt University[1], State University of New York at Albany[2],

University of Delaware[3], Carnegie Mellon University[4], Instructional Research Group,[5]

Metropolitan-Nashville Public Schools[6]

Inquiries should be sent to Lynn S. Fuchs, 228 Peabody, Vanderbilt University, Nashville, TN 37203.

This research was supported in part by Grant R324C100004 from the Institute of Education Sciences in the U.S. Department of Education to the University of Delaware, with a subcontract to Vanderbilt University, and by Core Grant #HD15052 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development to Vanderbilt University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Institute of Education Sciences, the U.S. Department of Education, the Eunice Kennedy Shriver National Institute of Child Health and Human Development, or the National Institutes of Health.

# Supported Self-Explaining during Fraction Intervention

Re-submitted June 26, 2015

**Abstract**

The main purposes of this study were to test the effects of teaching at-risk 4th graders to provide explanations for their mathematics work and examine whether those effects occur by compensating for limitations in cognitive processes. We randomly assigned 212 children to 3 conditions: a control group and 2 variants of a multi-component fraction intervention. Both intervention conditions included 36 sessions, each lasting 35 min. All but 7 min of each session were identical. In the 7-min component, students were taught to provide high quality explanations when comparing fraction magnitudes or to solve fraction word problems. Children were pretested on cognitive variables and pre/posttested on fraction knowledge. On accuracy of magnitude comparisons and quality of explanations, children who received the explaining intervention outperformed those in the word-problem condition. On word problems, children who received the word-problem intervention outperformed those in the explaining condition. Moderator analyses indicated that the explaining intervention was more effective for students with weaker working memory, while the word-problem intervention was more effective for students with stronger reasoning ability.

Key words: supported self-explaining, fractions, intervention, moderator, working memory, reasoning

**Supported Self-Explaining during Fraction Intervention**

Competence with fractions is important for advanced mathematics learning and success in the American workforce (National Mathematics Advisory Panel [NMAP], 2008; Geary Hoard, Nugent, & Bailey, 2012; Siegler et al., 2012). Yet understanding about fractions and skill in operating with fractions are difficult for many students (National Council of Teachers of Mathematics, 2007; NMAP; Ni, 2001). The NMAP therefore assigned high priority to improving fraction instruction. The focus of the present study was on the effects of intervention to enhance at-risk learners' performance on fractions at fourth grade. Our main purposes were to isolate the effects of intervention focused on teaching children to explain why fractions differ in magnitude and to examine whether effects accrue by compensating for limitations in cognitive processes.

In this introduction, we describe why teaching children to provide high quality explanations may enhance understanding of the explained content and provide a rationale for the supported self-explaining approach we took. We also discuss how self-explaining without such scaffolding can be cognitively demanding and why the approach we took may compensate for limitations in the cognitive resources many children with histories of poor mathematics learning experience. Finally, we describe the theoretical orientation of the multi-component intervention program in which our explaining intervention component was embedded.

**Explaining Why Fractions Differ in Magnitude**

A major purpose of the present study was to isolate the effects of intervention teaching children to provide sound explanations regarding a critical indicator of fraction understanding: comparing fraction magnitudes. Three types of self-explaining are described in the literature. *Spontaneous self-explaining* occurs when learners generate explanations for the to-be-learned material without being prompted to do so. A large literature demonstrates that individuals who spontaneously engage in self-explaining experience superior learning (e.g., Chi & Van Lehn, 1991; Siegler, 2002). However, not all learners spontaneously engage.

With *elicited self-explaining*, learners are prompted to invent explanations. Enhanced learning with elicited self-explaining can occur across a range of ages and domains, but results are mixed. In the most pertinent experiment, Rittle-Johnson (2006) provided insight for the inconsistency in findings. In one instructional session, she focused on whole-number mathematical equivalency problems with typically-achieving third through fifth graders. Prompting learners to self-explain promoted procedural accuracy more than a no-explanation condition. Yet, self-explaining did not produce more sophisticated procedures or understanding, perhaps because children's self-explanations rarely included a conceptual focus (even though students had been instructed to explain why solutions work). Instead, they tended to state whether an answer was correct or describe procedures to obtain answers. Their self-explanations also sometimes led to incorrect procedures. It therefore appears that inventing sound explanations is challenging and may depend on cognitive processes associated with strong learning, such as reasoning, working memory, and language comprehension. These findings also suggest that the key ingredient in self-explaining may be processing and expressing high-quality explanations, not inventing explanations. This is in line with Crowley and Siegler (1999) and Rittle-Johnson et al. (2015).

This brings us to the third form of self-explaining: *supported self-explaining*, in which learners operate on high-quality explanations already created for them. Because our target population was students with a history of poor mathematics achievement, many of whom experience limitations in the cognitive processes associated with mathematics learning and may therefore be especially vulnerable to inventing subpar explanations, the focus of the present study's explaining condition was supported explaining rather than invented self-explaining. We modeled high-quality explanations and provided children with practice in analyzing and applying the explanations, as we encouraged them to elaborate on and discuss important features of the explanations. We identified no previous randomized control trial testing the effects of this form of supported explaining among at-risk students, even though explaining is a broadly recommended instructional strategy and a strong focus in the mathematics Common Core State Standards (CCSS, 2013).

Based on the generative model of learning (Wittrock, 1990), which posits that learning requires actively processing and elaborating on the to-be-learned information while meaningfully connecting new information to existing knowledge, we hypothesized more accurate fraction magnitude comparisons for the explaining condition, compared to a contrast condition that received the same multi-component fraction program without the explaining component. This ensured that the contrast condition had high-quality relevant intervention on the same content. To control for intervention time, the contrast condition received a previously validated intervention component focused on fraction word problems requiring multiplicative reasoning.

We chose word problems as the focus of the contrast intervention for three reasons. First, accuracy in solving word problems strongly predicts employment and wages in adulthood (Bynner & Parsons, 1997). Second, multiplicative thinking, a major stumbling block for many students, is central to fraction knowledge, as reflected in the fact that finding equivalent fractions requires multiplying or dividing the numerator and denominator by the same quantity. Third, magnitude comparisons and word problems are both conceptually challenging, and the cognitive processes involved in fraction magnitude comparisons appear similar to those that support word problems. The word-problem intervention was based on schema theory, in which children are taught to conceptualize word problems as problem types (schemas); once they have identified the word-problem type, they apply the solution strategy for that problem type. Based on Fuchs et al. (in press), who showed positive effects for this word-problem intervention, we expected stronger word-problem solutions for the word-problem condition over the explaining condition.

**Does Supported Self-Explaining Compensate for Limitations in Cognitive Processes?**

Both fraction magnitude comparisons and word problems transparently demand reasoning ability. They also tax working memory in that (a) fraction magnitude comparisons require children to store and access information across a series of steps, which include finding equivalent fractions that need to be compared to benchmark fractions, while (b) word problems require children to process text describing a series of quantities, which must be sequentially evaluated and iteratively considered to create a coherent problem-solving model. Also, both tasks involve language comprehension because teachers use language to convey new ideas/procedures, because students use language when self-explaining, and because word problems are presented linguistically. In longitudinal studies aimed at predicting children's development of fraction and word-problem knowledge, roles have been demonstrated for reasoning (Fuchs et al., 2013; Seethaler et al., 2011), working memory (Fuchs, Schumacher, et al., 2013, 2014; Hansen et al., 2015; Jordan et al., 2013; Seethaler et al., 2011), and language comprehension (Fuchs et al., 2013; Jordan et al., 2013).

The present study extends this descriptive prediction literature by asking whether the effects of supported self-explaining occur, at least in part, by compensating for limitations in cognitive processes. To address this question, we assessed whether individual differences in reasoning, working memory, or language comprehension moderate the effect of the explaining versus word-problem condition on fraction magnitude comparisons (when all other variables in the model, such as performance on other cognitive processes and pretest performance, are held constant). We hypothesized that a *compensatory moderator effect* occurs in the following way. In the word-problem condition, we expected student outcomes to correlate with cognitive processing scores, such that students with severe cognitive deficits experience poorer outcomes than those with more adequate cognitive processing. On the other hand, we expected the supported explaining intervention to compensate for cognitive limitations, such that students achieve similarly regardless of where they fall along the cognitive processing distribution.

Finding such a compensatory moderator effect would provide further evidence for the role of these cognitive processes in self-explaining and would indicate that supported self-explaining increases learning by decreasing cognitive load. Also, because moderator analyses can specify where along the distribution of cognitive process scores the effect of supported self-explaining transitions from significance to nonsignificance, results may help schools identify which learners are more and less likely to benefit from the intervention. (We ran a parallel set of moderator analyses on the word-problem outcome, a point we return to in the discussion.)

In the most pertinent previous study, Fuchs et al. (2014) hypothesized a compensatory moderator effect for speeded strategic practice on the subtasks involved evaluating fraction magnitudes fraction understanding. This hypothesis was based on the assumption that creating automaticity on these subtasks would decrease cognitive load. Contrary to expectations, however, a disordinal moderator effect was identified: Students with severe working memory deficits profited more from practice in consolidating fraction ideas; those with more adequate working memory benefited more from the speeded strategic practice. Therefore, instead of a compensatory mechanism, results indicated that to benefit from speeded strategic practice, adequate working memory capacity is required.

The present study extends Fuchs et al. in ways that increase the potential meaningfulness of compensatory moderator effects. First, the present study’s intervention of major interest, supported self-explaining, was specifically designed to compensate for the cognitive resources required to engage in and derive benefits from high-quality explanations. Second, to extend control beyond Fuchs et al. (2014), the cognitive processes targeted in the present study were similarly relevant for both contrasted conditions. Third, self-explaining is broadly recommended in current education reform efforts.

**Context for the Two Fraction Intervention Components**

As mentioned, both intervention components were contextualized in a larger, multi-component intervention focused mainly on the measurement interpretation of fractions, which reflects cardinal size (Hecht et al., 2003) and is often represented with number lines (Siegler, Thompson, & Schneider, 2011). Such understanding can be linked to children's experiences with measuring, but depends largely on formal instruction that explicates the conventions of symbolic fraction notation (e.g., what the 3 and 4 mean in ¾), the inversion property of fractions (e.g., fractions with the same numerator become smaller as denominators increase), and the infinite density of fractions on any segment of the number line. Comparing fraction magnitudes is a central task within the measurement interpretation and, as such, both intervention conditions received a considerable amount of relevant instruction – one condition with the self-explaining component and the other condition without the self-explaining component.

The other form of understanding fractions relevant at fourth grade is the part-whole interpretation. This involves understanding a fraction as one or more equal parts of a single object (e.g., two of eight equal parts of a cake) or a subset of a group of objects (e.g., two of eight cakes). Such understanding is typically represented using an area model, in which a region of a shape or a subset of objects is shaded. It is intuitive, based on children's experiences with sharing, and is apparent in young children (Mix, Levine, & Huttenlocher, 1999).

The NMAP (2008) hypothesized that improvement in the measurement interpretation is more critical than the part-whole interpretation to fraction understanding. Fuchs, Schumacher, et al. (2013) provided support for this idea when they found that improvement in the measurement (not part-whole) interpretation mediates intervention effects on released fraction items from the National Assessment of Educational Progress (NAEP). Siegler and Ramani (2009) showed that a linear game board maps more effectively than a circular game board to the linear mental representation of numerical magnitude. Also, the part-whole interpretation encourages separate counting of numerator and denominator segments, increasing children's tendency to conceptualize a fraction as two separate whole numbers. By contrast, the measurement interpretation encourages relational thinking about numerators and denominators as determinants of a single number.

Even so, the part-whole interpretation continues to dominate American schooling (e.g., Fuchs et al., in press). Therefore, a secondary purpose of the present study was to replicate the efficacy of the multi-component fraction intervention, which focuses on the measurement interpretation in contrast to fraction instruction that focuses on the part-whole interpretation. This also permitted us to assess whether the explaining and word-problem conditions improved performance over what might be expected from the schools' typical fraction program. Therefore, we randomly assigned students to three conditions: two variants of intervention (with the explaining component vs. the word-problem component) and a control group focused on part-whole understanding. In line with previous studies and the NMAP, we hypothesized that effects would favor the two intervention conditions over the control group.

**Summary of Present Study's Purpose and Hypotheses**

To review, a major purpose of the present study was to isolate the effects of intervention that teaches children to provide sound explanations regarding a critical indicator of fraction understanding: comparing fraction magnitudes. To provide a stringent test of the effect of supported self-explaining, we contrasted fraction intervention with the self-explaining component against the same fraction intervention without the self-explaining component. To control for instructional time, the contrast condition included a validated word-problem component. At the same time, we examined whether the effects of supported self-explaining occur, at least in part, by compensating for limitations in cognitive processes. We hypothesized *compensatory moderator effect* in which student outcomes on fraction magnitude comparisons in the word-problem condition (which did not receive self-explaining) correlate with cognitive processing scores, but students in the supported self-explaining condition perform similarly regardless of where they fall along the cognitive processing distribution.

# Method

## Participants

All participants were fourth graders. We defined risk as performance below the $35^{th}$ percentile on a broad-based calculations assessment (Wide Range Achievement Test–4 [WRAT]; Wilkinson & Robertson, 2006). To ensure strong representation across the range of scores below the $35^{th}$ percentile, we sampled half the students from below the $15^{th}$ percentile and half from between the $15^{th}$ and $34^{th}$ percentiles. Because this study was about risk for mathematics difficulty, not intellectual disability, we administered the 2-subtest Wechsler Abbreviated Scales of Intelligence (WASI; Wechsler, 1999) to all students who met the risk criterion and excluded 15 children with T-scores below the $9^{th}$ percentile on both subtests. This permitted children with uneven cognitive profiles (as in learning disabilities) to remain in the study. We sampled 2-8 students per classroom, stratifying by more versus less severe risk in each classroom.

After exclusion and random sampling, the sample comprised 236 students from 52 classrooms in 14 schools. (Teachers in these classrooms were not aware of the study's hypotheses.) We randomly assigned these 236 students at the individual level, stratifying by classroom and risk severity, to the three conditions: intervention with the explanation component (EXP; $n = 79$), intervention with the word-problem component (WP; $n = 79$), and control (CON; $n = 78$). Six EXP, 5 WP, and 7 CON children moved before the end of the study. They did not differ statistically from the remaining students on pretest measures and did not differ significantly on any pretest measure as a function of condition. We omitted these children, leaving 218 students in the final sample: 73 in EXP, 74 in WP, and 71 in CON.

There were no significant differences among conditions on WRAT, with mean standard scores of 85.03 ($SD = 7.22$) for EXP, 85.61 ($SD = 7.76$) for WP, and 85.41 ($SD = 7.56$) for CON, or on WASI IQ, with mean standard scores of 93.63 (*SD* = 10.42) for EXP, 93.33 (*SD* = 12.57) for WP, and 93.27 (*SD* = 11.40) for CON. In the three groups, respectively, 58, 55, and 48% were female; 15, 17, and 19% were English learners; 93, 90, and 87% received subsidized lunch; and 7, 13, and 13% received special education. In EXP, the percentages of African-American, white, Hispanic (all white), and other students was 49, 19, 27, and 4; in WP, 57, 19, 23, and 1; in CON, 43, 21, 30, and 6. Chi-square tests indicated groups were demographically comparable.

**Screening Measures**

The mathematics screening measure was *WRAT-4-Math Calculations* (Wilkinson & Robertson, 2006) in which students complete calculation problems of increasing difficulty. Alpha on this sample was .87. The IQ screening measure was the WASI (Weschler, 1999). With *Vocabulary,* students identify pictures and define words. With *Matrix Reasoning*, students select the option that best completes a visual pattern. Reliability exceeds .92.

**Fraction Measures**

We assessed five outcomes. The first two pertained to these questions: Does EXP improve the accuracy of fraction magnitude comparisons?, Does WP enhance word-problem solving?, and Do cognitive processes moderate intervention effects between the two conditions? With the magnitude comparison/explanation measure, children identify which of two fractions is greater and explain why. The word-problem measure assesses skill with the types of multiplicative word problems addressed in that intervention component. The remaining three measures pertained to the replication of the multi-component fraction intervention. The generalized measure of fraction knowledge comprised released fraction items from the NAEP, which assess the part-whole and measurement interpretations of fractions with similar emphasis and are similarly distal across conditions (few items were aligned with any instructional tasks). The second measure, Siegler et al.’s (2011) Fraction Number Line Task, was not used for instruction in any study condition, but it indexes the measurement interpretation of fractions, which received greater emphasis in both intervention conditions than in the control group. The content of the third task, adding/subtracting fractions, received greater focus in the control group.

The *Fraction Magnitude Comparisons/Explanations* task, which is from the *Fraction Battery*-revised (Schumacher et al., 2013), comprises nine items: three with the same numerator; three with the same denominator; and three with a different numerator and denominator. Students place the greater than or less than symbol between fractions and write words/draw pictures to explain why the fractions differ in magnitude. The measure is scored as comparison accuracy (total number of items with the correct sign; max score = 9) and explanation quality (max score = 27). The focus of the explanation quality scoring is the conceptual content of explanations. For each items, students receive 1 point for explaining that the numerator indicates the number of parts and 1 point for explaining the denominator indicates the size of the parts. ("Both fractions have the same number of parts, but fifths are bigger than eighths so 4/5 is greater than 4/8" earns 2 points.) Students earn 1 more point if a picture shows two units of the same size, each showing the unit divided into the correct number of equal parts and the correct number of parts shaded. Alpha on this sample was .74 for accuracy and .92 for quality. Two independent coders scored all of the tests. On accuracy of magnitude comparisons, agreement was 100%; on explanation quality, computed point-by-point, 99.1%. Discrepancies were discussed and resolved.

*Multiplicative Word Problems*, from the *Fraction Battery-2012*-revised (Schumacher et al., 2013), includes six problems requiring students to make fractions from units (the "splitting" problem type), six problems requiring students to make units from fractions (the "grouping" problem type), and two distractor problems requiring students to compare fraction quantities (e.g., "Ruby ate ¼ of the pizza, and Bob ate 1/8 of the pizza. Who ate less pizza?"). None of the tested problems was used for instruction. Two near-transfer splitting problems rely on the vocabulary/question structure used in instruction (e.g., Lauren has 3 yards of ribbon. She cuts each yard of ribbon into sixths. How many pieces of ribbon does Lauren have now?). Four far-transfer splitting problems include novel vocabulary and/or questions (e.g., Jamie has 5 cups of batter to make cupcakes. Each cupcake needs 1/2 cup of batter. How many cupcakes can Jamie make? [novel vocabulary and question because *5 cups of batter* is the unit and *cupcakes* are the "pieces" in this problem, when students generally think of *cupcakes* as a unit]). Four near-transfer grouping problems incorporate unit fractions (e.g., Dante is making 8 peanut butter bars. Each peanut butter bar needs 1/4 cup of peanut butter. How many cups of peanut butter does Dante need?); two far-transfer grouping problems include non-unit fractions (e.g., Gabby needs to read 3 chapters in her book. Each chapter takes 2/3 of an hour to read. How many hours does Gabby need to spend reading?). The tester reads each item aloud while students follow along on paper. Students can ask for one rereading. For each problem, students earn 1 point for the correct numerical answer and 1 point for the correct label (e.g., pieces of ribbon). The maximum score is 26 (for distractor problems, only 1 point can be earned, which is for finding the correct numerical answer). Alpha on this sample was .82 - .90.

The *NAEP* measure comprised 19 released items from 1990-2009 *NAEP*: easy, medium, or hard fraction items from the fourth-grade assessment and easy items from the eighth-grade assessment. Testers read each problem aloud (with one rereading upon request). Eight items assess the part-whole interpretation (e.g., given a rectangle divided into six equal parts, the student shades 1/3); nine items assess the measurement interpretation (e.g., given four lists of three fractions, students identify which of the three fractions are arranged from least to greatest); one requires subtraction of fractions with like denominators; and one asks how many fourths make a whole. Students select an answer from four options (11 items); write answers (3 problems); shade a portion of a fraction (1 item); mark a number line (2 items); write a short explanation (1 item); or write numbers, shade fractions, and explain the answer (1 items with multiple parts). The maximum score is 25. Alpha on this sample was .71 - .78.

With the *Fraction Number Line* task (Hamlett, Schumacher, & Fuchs, 2011, adapted from Siegler et al., 2011), students place proper fractions, improper fractions, and mixed numbers on a number line labeled with endpoints of 0 and 2. A fraction is presented in large font below the number line. Students practice with two fractions and then estimate the location of 20 items: 12/13, 7/9, 5/6, ¼, 2/3, ½, 1/19, 3/8, 7/4, 3/2, 4/3, 7/6, 15/8, 1 1/8, 1 1/5, 1 5/6, 1 2/4, 1 11/12, 5/5, and 1. Items are presented in random order. The score for each item is the absolute difference between the placement and the correct position. Scores are averaged across items; divided by 2 (the numerical range of the number line); and multiplied by 100 to derive the percent of absolute error. Lower scores indicate stronger performance (in some analyses, we multiplied scores by -1). Test-retest reliability, on 63 students across 2 weeks, was .80.

From the *Fraction Battery*-revised (Schumacher et al., 2013), *Fraction Addition* includes five problems with like denominators and seven with unlike denominators; *Fraction Subtraction* includes six problems with like denominators and six with unlike denominators. In each subtest, half the problems are presented vertically and half horizontally. One point is awarded for the correct numerical answer; 2 points if reduced/rewritten one time (7 addition items; 8 subtraction items; e.g., ⅛ + ⅜ = 4/8 = ½; ¾ + 2/4 = 5/4 = 1 ¼); 3 points if reduced/rewritten two times (1 subtraction item: 10/6 - 2/6 = 8/6 = 1 2/6 = 1 ⅓). We used total score across subtests ($r$ = .83), with a maximum score of 41. Alpha on this sample was .90 - .94.

For NAEP, Calculations, and Word Problems, we had normative data on 265 not-at-risk students in 45 classrooms in 14 schools from a study conducted in the same district in the year preceding the present randomized control trial (Fuchs, Schumacher, et al., in press). We used these not-at-risk data to compute pre- and posttest achievement gaps for the at-risk conditions.

**Moderator Measures**

**Working memory**. To assess the central executive component of working memory, we used The *Working Memory Test Battery for Children* (WMTB-C; Pickering & Gathercole, 2001)-*Counting Recall*. The task includes six dual-task items at span levels from 1-6 to 1-9. Passing four items at a level moves the child to the next level. At each span level, the number of items to be remembered increases by one. Failing three items terminates the subtest. For each item, the child says how many objects are in an array and then recalls the series of counts in the trial. We used the trials correct score. Test-retest reliability ranges from .82-.91.

**Listening comprehension**. With *Woodcock Diagnostic Reading Battery* (*WDRB) - Listening Comprehension* (Woodcock, McGrew, & Mather, 2001), students supply the word missing at the end of sentences or passages that progress from simple verbal analogies and associations to discerning implications. At age 9, internal consistency reliability is .81.

**Reasoning**. *WASI Matrix Reasoning* (Wechsler, 1999) measures reasoning with pattern completion, classification, analogy, and serial reasoning tasks. Students select among five response options to complete a matrix, from which one section is missing. At age 9, internal consistency reliability is .94.

**Mathematics Instructional Time for the Two Interventions Conditions versus Control**

On questionnaires, teachers reported an average of 419.71 min of weekly mathematics instruction (*SD* = 81.67). They also reported that CON students received an average of 27.43 min (*SD* = 50.43) per week of supplemental math instruction, delivered mainly in small groups. This sums to ~447 min of mathematics instruction per week for CON students. Teachers also reported that the present study's intervention (105 min per week) typically occurred during part of classroom mathematics instruction or the school's intervention period. (When intervention did occur at other times, there was no systematic pattern to the types of instruction students missed.) Intervention students' weekly mathematics instructional time, including this study's intervention, averaged 453.63 min (*SD* = 57.61). So, on average, intervention students received 7 min per week more mathematics instruction compared to CON students (not a significant difference).

**District's Fraction Program**

The district's mathematics program was enVisionMATH (Scott Foresman-Addison Wesley, 2011). At fourth grade, the fractions units are "Understanding Fractions" and "Adding and Subtracting Fractions," with 70% of lessons allocated to understanding fractions. For understanding fractions, the program relies mainly on part-whole understanding by using shaded regions and other area model manipulatives, while encouraging students to write and draw when explaining fraction concepts. In a single lesson, benchmark fractions and equivalent fractions are introduced to teach students to make magnitude decisions (number lines are not included). Adding and Subtracting Fractions are taught via procedural rules. Fraction word problems are addressed by focusing mostly on additive reasoning and equal sharing (e.g., Eight friends divide 3 pizzas equally. How much does each friend eat?), with a smaller emphasis on multiplicative reasoning (e.g., Danielle bought 3 ¼ yards of ribbon. How many pieces of ¼-yards of ribbon did Danielle buy?). As per What Works Clearinghouse (WWC; U.S. Department of Education, 2013), enVisionMATH has "potentially positive effects." The one study meeting WWC standards (Resendez & Azin, 2008) reported an average effect size (ES) of 0.15 at grade 4. Only 1.9% of teachers reported relying exclusively on this program. Most (84.6%) relied on a combination of enVisionMATH and CCSS, while 13.5% relied exclusively on CCSS.

**Distinctions among the Three Study Conditions**

See Table 1 for teacher questionnaire responses concerning CON fraction instruction, which includes the schools' intervention period, as contrasted to the EXP and WP intervention conditions. Note that instructional practice information was reported at the classroom level and the amount of intervention time for each at-risk student was reported at the individual student level. (Classroom instruction was not recorded or directly observed.)

To describe the relative emphasis on different types of fraction representations, teachers distributed 100 points across response options. Part-whole representations (tiles, circles, pictures with shaded regions, blocks) constituted 75.96% of their emphasis; number lines accounted for

20.87%. By contrast, number lines reflected 70% of the intervention's emphasis, with only 30% on part-whole interpretation.

For instructional strategies to help students identify the relative magnitude of fractions, CON relied heavily on finding common denominators, rather than focusing on understanding why fractions have differing magnitudes. When combined with cross multiplying (addressed procedurally), 41.82% of the CON emphasis was procedural – not focused on the ideas essential to why fractions magnitudes differ. Yet, 56.15% of the CON emphasis was on activities and strategies that address ideas about why fraction magnitudes differ: thinking about relative placement on number lines, comparing fractions to benchmark fractions, using manipulatives, and considering the meaning of numerators and denominators. (As shown in Table 2, the remaining 2.03% of responses were "other.") By contrast, 90% of intervention activities focused on these conceptual strategies/activities.

CON students explained their work with fractions an average of 4.36 times per week (*SD* = 1.68), as reported by the teachers. This contrasts to an average of 21 times per week for the EXP condition and 15 times per week in the WP condition (EXP and WP provided explanations for fraction work in other portions of the multi-component program, not just in the EXP condition's 7-min component; see below). The CON group put more emphasis on drawing pictures than in EXP or WP, while EXP and WP emphasized providing explanations via words in oral and written form. EXP students were also taught to label pictures.

For word problems, CON included the following strategies to help students understand word-problem narratives: drawing pictures, writing equations, using words to explain thinking, and making tables. Yet, 20.13% of the instructional emphasis was on key words, which can discourage deep thinking about word problems and often produces incorrect solutions (Schumacher & Fuchs, 2012). The WP condition did not address key words and instead relied on using words to explain thinking, identifying problem types, and using arrays to represent the multiplicative structure of problems. Only WP (not EXP) received intervention on word problems; so, EXP students' word-problem instruction was the same as in the CON condition.

Thus, four major differences between CON and EXP/WP were as follows. (1) CON focused on part-whole understanding; EXP/WP emphasized measurement interpretation. (2) CON did not restrict the range of fractions; EXP/WP limited the range of denominators (2, 3, 4, 5, 6, 8, 10, 12) and equivalent fractions (1/2, 1/3, 2/3, ¼, 1/5, 2/5, 1/1). (3) CON required students to explain work substantially less often than EXP/WP, and CON explanations emphasized words more than pictures while the reverse was true in EXP/WP. (4) For word problems, CON focused more on drawing pictures, making tables, and using key words; WP focused more on words to explain thinking, identifying problems as belonging to problem types, and representing the structure of problems.

**The Multi-Component Fraction Intervention**

The multi-component fraction intervention is referred to as *Fraction Face-Off!* (see Fuchs, Schumacher, et al., in press, for tutor training, the structure of the manual/materials, and activities). Sessions occurred 3 times per week for 12 weeks with pairs of children. Tutors were employees of the research grant. Some were licensed teachers; most were not. Each was responsible for 2-4 groups, distributed across the EXP and WP conditions. To avoid contamination between conditions, we color coded materials, conducted periodic live observations, regularly monitored fidelity of implementation audiotapes, and provided guidance and engaged in problem solving during bi-weekly meetings.

Of each 35-min session, 28 min were the same in EXP/WP, primarily focused on the measurement interpretation of fractions, with instruction on comparing, ordering, placing fractions on number lines, and equivalencies. This was preceded by attention to the part-whole interpretation (showing objects with shaded regions) and equal sharing to build on prior knowledge and classroom instruction. Number lines, fraction tiles, and fraction circles were used to introduce and review concepts across the 36 lessons. See Figure 1 for timeline and see Fuchs, Schumacher, et al. (in press) for additional information.

Instruction was explicit. Tutors (a) introduced topics with worked examples, while explaining how work is completed and what the tutor thought when completing each step; (b) provided explanations in simple, direct language, with students explaining their thinking in their own words; (c) relied on efficient solution strategies and required students to practice applying strategies while they explained how/why strategies make sense; (d) ensured students had the necessary background knowledge/skills to succeed with strategies; (e) provided practice so students used strategies to generate many correct responses; and (f) included systematic cumulative review.

To address at-risk students' difficulties with attention, motivation, and self-regulation (Montague, 2007), tutors taught the meaning of *on-task behavior* and its importance for learning. A timer beeped at 3 unpredictable times per lesson. If all students were on task at the beep, each earned a checkmark. In The Individual Contest, two problems were pre-designated bonus problems for which students received checkmarks if their work was completed correctly, but students were not told which problems earned a bonus point until all work was completed. Tutors awarded a "half dollar" or "quarter dollar" for each checkmark. Once per week, tutors opened the "Fraction Store," where students spent earnings on small prizes priced at $1, $7, $13, or $20, exchanging fraction dollars for whole dollars to determine which prizes they could afford or saved for more expensive prizes.

**EXP and WP Components**

With <u>EXP instruction</u> (see Figure 1 for timeline), tutors modeled a 4-step problem-solving sequence, gradually transferring responsibility to students while scaffolding to ensure understanding. In the first step, students wrote whether the fractions had the same denominator ("same D"), the same numerator ("same N"), or different numerators and denominators ("both diff”). In the second step, students explained why a good drawing shows two units of the same size, with each unit divided into the correct number of parts, with all parts in each unit the same size, and with the correct number of parts shaded. In the third step, students labeled each drawing of a fraction with its numerical value and described how the parts in the fractions compared (“same size parts” to indicate each unit was divided into the same number of parts or “bigger parts” to indicate the fraction with fewer parts had bigger parts). When numerators and denominators both differed, students rewrote ½ so both fractions had the same denominator and updated the picture to show how ½ is equivalent to the original fraction and how finding a common denominator allowed them to compare two fractions with the same size parts. Then students circled the larger fraction. In the final step, students wrote a short sentence or phrase to explain why the circled fraction was greater (e.g.., “more same size parts means bigger fraction”; “same number of parts, but fourths are bigger than eighths”). Tutors provided corrective feedback as needed. The ideas underlying EXP instruction were the same as those addressed in the multi-component intervention. Later lessons focused on discriminating between viable explanations for same denominator versus same numerator problems. Tutors solved a same denominator and same numerator problem side by side. Students highlighted and discussed important distinctions between the problem types. Note, however, that we did not incorporate schema-based instruction into self-explaining (i.e., students did not classify problems into problem types prior to formulating explanations).

WP instruction (see Figure 1 for timeline) relied on schema theory, with which students categorize WPs as belonging to a WP type based on their underlying mathematical structure and then apply a WP-solving strategy specific to that WP type. Once both WP types were taught, practice included distractor WPs in which students identified the larger or smaller fraction. The goal was to increase students’ ability to recognize non-examples of the taught WP types, decrease the tendency to overgeneralize strategies, and add practice on comparing fractions.

The two WP types taught were “Splitting” and “Grouping.” Tutors introduced Splitting WPs first, using an intact story about a “Splitting” situation (e.g., Melissa had 2 lemons. She cut them in half. Then she had 4 pieces). Tutors used fraction circles to explain what cutting 2 lemons to produce 4 pieces means. Then, they presented the same story, with the final sentence as a question, “How many pieces of lemon did she have?” Tutors completed a worked example, providing a rationale for each step of the solution strategy and using fraction circles to explain the solution methods. Then tutors taught that we call WPs with the same structure as the “Melissa” problem “Splitting” WPs, which describe a unit being cut, divided, or split into equal parts. On the next day, tutors reviewed the underlying structure of the Splitting WP type. Then, they presented a novel Splitting problem, and students explained why it was a Splitting problem. Next tutors taught a series of steps to help students organize their papers, synthesize information in the WP, and solve the WP: (1) underline the unknown; (2) identify the units and the size of each piece and label them “U” and “S”; (3) draw an array to represent the structure of splitting WPs; (4) enter into the array information from the WP to show each unit divided into the fractional pieces (e.g., for each unit divided into fifths, they wrote 1/5 to represent each piece for each unit); and (5) solve the WP and write the numerical answer and word label.

To challenge and extend identification of Splitting WPs and encourage flexibility in WP-solving strategies, we incorporated transfer features that altered the language typically found in splitting WPs. This included activities to promote recognition of different vocabulary (e.g., synonyms for *pieces* such as *wedges*) and promote recognition of unfamiliar question formats. Consider this problem: The relay race is 4 miles. Each leg of the race is ½ mile. How many kids do we need to run the relay race? This example has novel vocabulary and a novel question because there are no familiar vocabulary words initiating a Splitting or dividing action and because *kids*, the “pieces” in this WP, are not typically thought of as pieces.

"Grouping" problems (e.g., Keisha wants to make 8 necklaces for friend. For each necklace, she needs ½ of a yard of string. How many yards of string does Keisha need?) were introduced using parallel methods, with two major distinctions that clarified the underlying structure of Grouping problems. First, students identified "items" (instead of units), which refers to how many fractional pieces are needed. Second, the array was structured to accommodate WP-solving strategies representing this underlying structure. Later grouping problems included non-unit fractions, which required an additional WP-solving step.

**Fidelity of Implementing Intervention**

Every intervention session was audiotaped. We randomly sampled 20% of 2,556 recordings to represent tutor, student, and lesson comparably. A research assistant listened to each sampled tape, while completing a checklist to identify essential points implemented. The mean percentage of points addressed was 98.43 (*SD* = 2.03) in EXP and 98.57 (*SD* = 4.66) in WP. Two research assistants independently listened to 20% of the 511 coded recordings to assess concordance. The mean difference in score was 1.27%. With a test for dependent samples, we assessed the comparability of fidelity between conditions because each of the 19 tutors conducted sessions in both conditions. The difference was not significant, $t(18) = 0.13$, $p = .896$.

**Procedure**

In August/September, testers *screened* in large groups using WRAT; then administered the WASI individually to students who met the WRAT risk criterion. In September/October, to assess *pretreatment comparability among study groups on fraction knowledge*, testers administered the NAEP, the Fraction Addition/Subtraction measure, the magnitude comparison/explanation measure, and the word-problem measure in two small-group sessions. The Fraction Number Line task was administered individually. From late October to early February, *intervention* occurred for 12 weeks. In early March (after an intervening spring vacation), we assessed *intervention effects*, by re-administering Fraction Addition/Subtraction, NAEP, the magnitude comparison/explanation measure, and the word-problem measure in two small-group sessions and re-administering Fraction Number Line individually. Time permitted students to complete all tests. All sessions were audiotaped; 20% of tapes were randomly selected, stratifying by tester, for accuracy checks by an independent scorer. Agreement on test administration and scoring exceeded 98%. Testers were blind to conditions when administering and scoring tests.

## Results

### Preliminary Analyses

We conducted preliminary analyses to evaluate the nested structure of the data (a cross-classified partially nested design in which nesting occurred at the classroom level for all three conditions and at the tutoring-group level for the two intervention conditions). First, we estimated the proportion of variance in each outcome due to classrooms and tutoring groups. Respective intraclass correlations were negligible to small (~.00 and ~.00 for number line; .03 and .01 for calculations; .01 and .05 for NAEP; ~.00 and ~.00 for EXP, and ~.00 and .02 for WPs). Then, we examined the random effects model for EXP and WP, while accounting for nesting involved in the tutoring conditions. Results indicated the random effects due to tutoring clusters could be ignored. Next, having dropped the tutoring clusters and addressed all three conditions together, we ran multilevel regression models examining tutoring effects on each outcome while accounting for nesting at the classroom level. Results did not alter any inferences based on single-level models. Given these results along with the fact that random assignment occurred at the individual student level, we report single-level analyses.

We then confirmed that pretest performance of the three groups on each fraction measure was comparable. Next, because we relied on a residualized change approach to analyze effects of condition (covarying pretest performance to reduce within-group error variance), we assessed the homogeneity of regression assumption, which was met for all measures except word problems,

$F(2,208) = 23.43$, $p < .001$. Therefore, in models involving word problems, we controlled for the interaction between pretest word-problem score and condition.

**Does Teaching Students to Explain Their Mathematics Work Enhance Accuracy of Magnitude Comparisons and Quality of Explanations?**

We conducted 1-way analyses of variance (treatment condition was the factor) on comparison accuracy and explanation quality posttest scores (only posttest data were collected). Although our question centered on the contrast between the EXP versus WP conditions, we included the CON group to assess whether the EXP or WP condition outperformed CON. (Given the multi-component intervention's focus on the measurement interpretation, for which magnitude comparisons is a central task, we expected WP to outperform CON.) Table 2 shows posttest means and *SD*s by condition. Table 3 shows inferential test statistics, pairwise follow-up comparisons (Fisher least significant difference post hoc procedure; Seaman, Levin, & Serlin, l991), and ESs comparing conditions (difference in posttest scores divided by the pooled posttest *SD*). On comparison accuracy, EXP outscored WP (ES = 0.43), and both intervention conditions outperformed CON (ES = 1.37 for EXP vs. control; 0.89 for WP vs. control). On explanation quality, EXP provided higher quality explanations than WP (ES = 0.93) and CON (ES = 1.37), and WP provided higher quality explanations than CON (ES = 0.60). (We also ran an analogous analysis of variance controlling for language comprehension. Results were similar, $F(2,208)=32.98$, $p < .001$, with respective ESs of 0.98, 1.29, and .54.)

**Does WP Enhance Word-Problem Solving?**

We ran analysis of covariance on the word-problem outcome to assess the efficacy of the WP component's efficacy. The model included the pretest covariate and the interaction between pretest word-problem score and condition. Again, we included the CON group to assess whether the WP or EXP condition outperformed CON. (Because neither EXP nor the multi-component intervention included word problems, we did not expect an advantage for EXP.) Table 2 shows pretest, posttest, and posttest means adjusted for pretest scores by condition. Table 3 shows inferential test statistics, follow-up tests (Fisher least significant difference post hoc procedure), and ESs comparing conditions (differences in adjusted posttest scores divided by the pooled posttest *SD*). WP outscored EXP (ES = 1.48) and CON (ES = 1.20), with no significant difference between EXP and CON (ES = 0.12). For WPs, normative data for not-at-risk peers were available to estimate pre- and posttest achievement gaps for at-risk study participants by condition (see Table 2, last 3 columns). These gaps are expressed as ESs (raw score difference in means, divided by the not-at-risk *SD*). The pretest gap on fraction word problems was ~0.90 for each condition. At posttest, the gap closed for WP, but increased somewhat for EXP and CON.

**Do Individual Differences in Cognitive Processes Moderate Effects between EXP and WP?**

To test whether a moderator effect qualified the main effects between the EXP and WP conditions, we followed the recommendation of Preacher and Hayes (2008; Hayes, 2012), using an ordinary least squares path analytic framework to evaluate the moderator effects of pretest reasoning, working memory, and language comprehension. We conducted three analyses on the comparison accuracy score. In each analysis, one of the three cognitive variables was tested as the moderator, while controlling for the other two cognitive/linguistic variables. We ran three analogous models for the word-problem outcome, while also controlling for pretest word-problem skill and the interaction between pretest word-problem score and condition. Sample-based standard scores across the two conditions were employed for all measures. See Table 4 for raw and standard score means/*SD*s and correlations for variables included in these analyses. See Table 5 for results of the moderation analyses.

For analyses conducted on comparison accuracy, $R^2 = .20$ for working memory as the tested moderator, $F(5,136) = 6.81$, $p < .001$; $R^2 = .18$ for language comprehension as the tested moderator, $F(5,136) = 6.02$, $p < .001$; and $R^2 = .18$ for reasoning as the tested moderator, $F(5,136) = 6.07$, $p < .001$. Working memory was the only significant moderator (see Table 5 for interaction *F* values).

To probe this interaction between working memory and the effect of EXP over WP, we relied on the Johnson-Neyman technique (Bauer & Curran, 2005; Hayes & Matthes, 2009), which derives the value along the full continuum of the moderator at which the effect of X on Y transitions from statistical significance to nonsignificance. Figure 2 is a visualization of the significant interaction, with illustrative points on the distribution of *within*-sample working memory percentile ranks (x-axis). As shown, for students with lower working memory, the EXP condition (white bars) produced significantly stronger outcomes than the WP condition (black bars, i.e., non-EXP). However, for students with more adequate working memory, effects favoring EXP were not significant. As shown, this is because the relation between working memory and the outcome was stronger in the WP condition than in the EXP condition. The effect favoring EXP over WP transitioned to nonsignificance at 1.09 *SD*s above the at-risk *sample* mean on working memory. Thus, consistent with expectations, a compensatory moderator effect was demonstrated.[1]

For the word-problem outcome, $R^2$ = .49 for working memory as the tested moderator, $F(7,134) = 18.33$, $p < .001$; $R^2$ = .48 for language comprehension as the tested moderator, $F(7,134) = 17.61$, $p < .001$; and $R^2$ = .51 for reasoning as the tested moderator, $F(7,134) = 20.20$, $p < .001$. Reasoning was the only significant moderator (see Table 5 for interaction *F* values). Figure 3 shows that effects favoring WP (now white bars) over EXP (now black bars, i.e., non-WP) were not significant for students with very low reasoning, but were significant for students with more adequate reasoning. This is because the relation between reasoning and the outcome was stronger in the WP condition than in the EXP condition. The effect favoring WP transitioned to nonsignificance at 1.81 *SD*s below the at-risk *sample* mean on reasoning. Thus, the intervention did not adequately compensate for the word-problem demands.

**Does the Multi-Component Intervention Enhance Other Forms of Fraction Knowledge More than the Control Group?**

We conducted 1-way analyses of covariance on NAEP, number line, and calculation posttest scores, controlling for pretest scores on the relevant measure. See Table 2 for pretest, posttest, and posttest means adjusted for pretest scores and for achievement gap data for NAEP and calculations, by condition. Table 3 shows inferential test statistics, follow-up tests (Fisher least significant difference post hoc procedure), and ESs between conditions (differences in adjusted posttest scores divided by the pooled posttest *SD*). NAEP, number line, and calculations performance was stronger for each intervention condition compared to CON. ESs favoring EXP over CON ranged from 0.57 to 1.98; ESs favoring WP over CON ranged from 0.70 to 2.08. Thus, each intervention condition outperformed CON on generalized fraction knowledge (NAEP), the measurement interpretation of fractions (number line), and fraction calculations.

(We also tested moderator effects between each intervention condition and the control group on magnitude comparisons and word problems. This was not central to our research questions, but can be found at *provide link here [this is at end of this file for review purposes]*.)

## Discussion

The main purpose of this study was to test the effects of teaching at-risk fourth graders to provide explanations for their mathematics work, while examining whether those effects occur by compensating for limitations in basic cognitive processes. We focused on children's explanations because high-stakes assessments increasingly emphasize student explanations of mathematics work. This emphasis reflects the belief that when children explain mathematics work, their understanding deepens (Kilpatrick et al., 2001; Whitenack & Yackel, 2002). We relied on a supported self-explaining approach, with which we modeled high-quality explanations and provided children with practice in analyzing and applying the explanations, as we encouraged them to elaborate on and discuss important features of the explanations.

### Does Supported Self-Explaining Enhance Content Knowledge and Explanation Quality?

In terms of effects on students' conceptual content knowledge – the accuracy with which they identified larger and smaller fractions, EXP outperformed the WP condition with a moderate ES of 0.43. On the quality of explanations about why fraction magnitudes differ (scored for the explanations' conceptual content), effects more dramatically favored EXP over WP, with an ES of 0.93. These outcomes for EXP over WP are noteworthy because the multi-component fraction intervention provided children in *both* conditions the same instruction on essential ideas and efficient procedural strategies for comparing fraction magnitudes. What distinguished EXP from WP was supported self-explaining.

Results provide support for a generative model of learning (Wittrock, 1990), which posits that learning requires actively processing and elaborating on the to-be-learned information while meaningfully connecting new information to existing knowledge. Results also corroborate studies showing that the benefits of self-explaining can be derived from practicing and operating on high quality explanations, without inventing those explanations, as in Crowley and Siegler (1999) and Rittle-Johnson et al. (2015). Although prior work on this point is inconsistent (Kwon, Kumalasari & Howland, 2011; Schworm & Renkl, 2002), the positive relation between explanation quality and learning is consistent across sources of evidence. So, it is concerning that even typically developing children experience difficulty inventing sound conceptual explanations (Rittle-Johnson, 2006). The ability to create such explanations may depend on the cognitive processes associated with strong learning. Thus, for students with histories of poor mathematics learning, who experience limitations in such basic cognitive processes and are at-risk for generating subpar explanations, supported self-explaining may provide a better option. Future research should directly contrast these two forms of self-explaining among at-risk learners.

**Does Supported Self Explaining Compensate for Limitations in Cognitive Processes?**

We tested the hypothesis that the effects of our supported self-explaining intervention occurred via such a compensatory mechanism. Toward this end, we assessed whether individual differences in each targeted cognitive ability (reasoning, working memory, language comprehension) moderated the effect favoring the EXP over the WP condition on fraction magnitude comparisons. We hypothesized what we refer to as a *compensatory moderator effect*. In the contrast condition (i.e., the WP condition that did not receive the EXP component), we expected student outcomes to correlate with cognitive processing scores, such that students with severe cognitive deficits experience poorer outcomes than those with more adequate cognitive processing. By contrast, we expected the supported explaining intervention to compensate for cognitive limitations, such that students achieve similarly regardless of where they fall along the cognitive processing continuum.

We found such a compensatory moderator effect for working memory, while controlling for the effects of reasoning and language comprehension. The accuracy of magnitude comparisons in the EXP condition (white bars in Figure 2) were similar regardless of students' working memory capacity (the correlation between working memory and magnitude comparisons in the EXP condition was -.04). By contrast, accuracy in the non-EXP condition (the WP condition, shown with black bars) significantly correlates with students' working memory ($r$ = .32). Without EXP, students with weaker working memory suffered lower accuracy than did students with more adequate working memory, and the effect favoring EXP over non-EXP transitioned from significance to nonsignificance at 1.09 *SD*s ($86^{th}$ percentile) above the sample's mean on working memory.

Finding this compensatory moderator effect extends the longitudinal, descriptive prediction literature by providing a more stringent form of evidence for working memory's role in mathematics learning generally and in processing fraction magnitude comparisons in particular. It also suggests a stronger role for working memory than for reasoning or language comprehension – at least in the context of a multi-component fraction intervention that provides explicit instruction on understanding and executing those comparisons. More generally, the significant compensatory moderator effect for working memory on fraction magnitude comparisons indicates that supported self-explaining increases learning by decreasing cognitive load. More practically, by revealing where along the distribution of cognitive process scores the effect of supported self-explaining transitions from significance to nonsignificance, results may help schools identify which learners are more and less likely to benefit from such intervention.

**Were Effects of the WP Component Also Moderated by Basic Cognitive Processes?**

On the word-problem outcome measure, the main effect favoring WP over EXP was strong, with an ES of 1.48. Our WP instructional approach was rooted in schema-based instruction, which is designed to compensate for limitations in students' cognitive resources by teaching them to identify word problems as belonging to categories that share structural features (splitting vs. grouping) and to represent the underlying structure of the word-problem category with a number sentence (Fuchs et al., 2003, 2009, 2010) or visual display (Fuchs, Schumacher, et al., in press; Jitendra & Star, 2012, Jitendra, Star, Rodriguez, et al., 2011; Jitendra, Star, Starosta, et al., 2009). The main effect favoring WP over EXP replicates the efficacy of our fraction word-problem intervention component (Fuchs, Schumacher, et al., in press). It also corroborates previous schema-based instruction research on word problems with whole and rational numbers (Fuchs et al., 2003, 2009, 2010; Jitendra & Star, 2012; Jitendra et al, 2011; Jitendra et al., 2009).

Even so, because schema-based instruction is designed to compensate for limitations in students' cognitive resources, we hypothesized a compensatory moderator effect involving WP. Although we found a significant moderator effect (while controlling for the effects of working memory, language comprehension, pretest word-problem performance, and the interaction between scores intervention), it did not indicate a compensatory mechanism. Instead, it suggested that responsiveness to the WP component grew stronger as reasoning ability increases. That is, word-problem outcomes for correlated with reasoning ability more strongly for students in the WP condition (white bars in Figure 3, $r$ = .36) than for students in the EXP (non-WP) condition (black bars in Figure 3, *r* = .05). The effect favoring WP over non-WP transitioned from significance to non-significance at 1.81 *SD*s (4$^{th}$ percentile) below the sample's mean on reasoning. This suggests that the WP intervention's instructional methods demanded higher quality reasoning than the EXP (non-WP) condition.

This makes sense for the following reason. The EXP intervention provided no word-problem instruction. Therefore, EXP students called upon strategies taught in their classrooms/school interventions to solve word problems. As in Table 1, these strategies required superficial analysis (e.g., key words) rather than deep reasoning. By contrast, in relying on the schema-based strategies, WP students exercised more analytical strategies. As such, students with more severe reasoning deficits benefited less from WP intervention than students with more adequate reasoning. This prerequisite ability moderator effect therefore reveals the need to develop the intervention further to address the full range of learners' reasoning ability.

Both types of moderator effects qualify main effects in ways that identify characteristics that make students more or less responsive to intervention and therefore can help schools make service delivery more efficient and effective. Results suggest that students with adequate working memory do not require the EXP intervention component but instead benefit more from the WP intervention component. By contrast, students with very low reasoning ability require additional support to benefit from the WP component. In this way, a prerequisite ability moderator effect also permits insight about the need to strengthen intervention so that the needs of students with very low cognitive or linguistic processes are addressed more effectively.

In this vein, results reveal a potential avenue for strengthening the WP condition for students who did not respond to intervention. The effect of language comprehension in the model testing the reasoning interaction on the word-problem outcome was large (2.62), while the corresponding effect on the explanation outcome was 0.07. This reflects the importance of language comprehension in multiplicative word problems. Future research should examine whether a focus on language comprehension in the WP intervention further decreases inadequate response to intervention.

We nevertheless remind readers that moderation analyses are correlational, so inferences about causation should be applied cautiously. Also, although these interactions suggest the need to account for individual differences in response to intervention, corroborating studies are required. Research is specifically needed to address the challenges of "personalizing" instruction, because many at-risk students experience limitations across multiple cognitive processes.

The hope, however, is that identifying such moderator effects may eventually help schools deploy interventions in ways that improve effectiveness and efficiency, and this study's moderator effects involving the contrast between WP and EXP illustrate how this might occur. At the same time, it is important to note that the pool of children who did not respond to the WP intervention was very small: The effect of the WP condition over the EXP (non-WP) condition transitioned to nonsignificance when reasoning ability fell below the 4th percentile of this at-risk sample's mean (and this sample scored one-half *SD* below the test's normative framework; see Table 4). Thus, the WP intervention component was effective, relative to the EXP condition, for all but a small group of learners with extremely low reasoning. Moreover, effects were robust for the contrast between the EXP condition and the *control* condition on the content knowledge that served as the focus of the self-explaining, with a large ES of 1.37, and for the contrast between the WP condition versus the *control* condition on word problems, with a large ES of 1.20.

**Effects of the Multi-Component Fraction Intervention**

This brings us to the overall effects of the multi-component fraction intervention. In the present randomized control trial, both versions of the multi-component fraction intervention, each focused mainly on the measurement interpretation of fractions, outperformed the control group, which focused mainly on the part-whole understanding of fractions. ESs ranged from 0.63 to

1.98 for the multi-component version that included the EXP component and from 0.71 to 2.08 for the multi-component version that included the WP component.

This advantage for both intervention conditions over the control group included the NAEP items, which assessed measurement and part-whole interpretations with comparable emphasis, as well as fraction calculations, even though the control group focused more than the intervention conditions on this topic. This supports previous findings suggesting that understanding of fractions supports accurate fraction procedures (Hecht et al., 2003; Mazzocco & Devlin, 2008; Ni & Zhou; 2005; Rittle-Johnson et al., 2001). It is also important to note that at this grade level, when classroom fraction instruction gains momentum, the control group’s achievement gap with respect to not-at-risk student local norms grew approximately 70% on NAEP and nearly doubled on calculations. By contrast, the achievement gap of at-risk students who received the multi-component intervention held approximately steady on NAEP and closed entirely on calculations.

In these ways, each intervention condition produced superior fraction learning for at-risk fourth graders compared to the business-as-usual at-risk control group. These results echo earlier randomized control trials suggesting the importance of supplemental explicit fraction intervention (Fuchs et al., 2015) that incorporates a strong emphasis on the measurement interpretation of fractions (Fuchs, Schumacher et al., 2013, 2014, in press). Finding that the measurement interpretation of fractions is a key instructional target is also in line with the perspective expressed in the NMAP report (2008).

Before closing, however, we caution readers about several study limitations. First, the control group received less instructional time than the two intervention conditions in groups of two students – although the difference in time was not statistically significant. Moreover, it is important to note that this limitation does not apply to the main study analyses, contrasting the EXP and WP conditions. These two conditions received the same amount of 2:1 intervention time. Second, in the explanation quality analyses, we did not control students' writing skill but as noted in the results section, controlling for language comprehension (which is related to written expression) did not alter results. Also, the writing demands on the explanation quality measure were not extensive. This caveats notwithstanding, we conclude that the overall efficacy of the multi-component fraction intervention and for its EXP and WP components appears strong. Future studies should extend this line of work by examining how instruction on decimals might strengthen fraction understanding and might employ longitudinal designs to assess maintenance and determine which children require longer-term fraction intervention for sustained success beyond fourth grade.

**Footnote**

[1] In Table 4, correlations between some moderator variables [measured at the start of the study] and outcomes [measured at the end of the study] are low because during the study, intervention was delivered to disturb the natural relation between the variables, but only for half the sample - for the condition that received relevant intervention. Across the two conditions, as shown in Table 4, this reduces the magnitude of these relations.

**References**

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: inferential and graphical techniques. *Multivariate Behavioral Research, 40*, 373-400. doi: 10.1207/s15327906mbr4003_5

Bynner, J., & Parsons, S. (1997). *Does Numeracy Matter?* London: The Basic Skills Agency.

Chi, M.T.H., & Van Lehn, K.A. (1991). The content of physics self-explanations. *Journal of the Learning Sciences, 1*, 69-105. doi:10.1207/s15327809jls0101_4

Crowley, K., Siegler, R.S. (1999). Explanation and generalization in young children's strategy learning. *Child Development, 70*, 304-316. doi: 10.1111/1467-8624.00023

Fuchs, L. S., Fuchs, D., Powell, S. R., Seethaler, P. M., Cirino, P. T., & Fletcher, J. M. (2008). Intensive intervention for students with mathematics disabilities: Seven principles for effective practice. *Learning Disability Quarterly, 31*, 79-92. doi: 10.2307/20528819

Fuchs, L.S., Fuchs, D., Compton, D.L., Wehby, J., Schumacher, R.F., Gersten, R., & Jordan, N.C. (in press). Inclusion versus specialized intervention for very low-performing students: What does *access* mean in an era of academic challenge? *Exceptional Children.*

Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., Hosp, M., Jancek, D. (2003). Explicitly teaching for transfer: Effects on third-grade students' mathematical problem solving. *Journal of Educational Psychology, 95*, 293-305. doi:10.1037/0022-0663.95.2.293

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., Hamlett, C. L., Zumeta, R. O. (2009). Remediating number combination and word problem deficits among students with mathematics difficulties: A randomized control trial. *Journal of Educational Psychology, 101*, 561–576. doi:10.1037/a0014701

Fuchs, L. S., & Schumacher, R.F. (2011). *Fraction Face-Off!*. Available from L.S. Fuchs, 228 Peabody, Vanderbilt University, Nashville, TN 37203.

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

Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Malone, A. S., Wang, A., Hamlett, C. L., Jordan, N. C., Gersten, R., Siegler, R. S., Changas, P. (in press). Effects of intervention to improve at-risk fourth graders' understanding, calculations, and word problems with fractions. *The Elementary School Journal.*

Fuchs, L. S., Schumacher, R. F., Sterba, S.K., Long, J. Namkung, J., Malone, A. S., Hamlet, C. L., Jordan, N. C., Gersten, R., Siegler, R. S., Changas, P. (2014). Does working memory moderate the effects of fraction intervention? An aptitude-treatment interaction. *Journal of Educational Psychology, 106,* 499-514. doi:10.1037/a0034341

Fuchs, L. S., Zumeta, R. O., Schumacher, R. F., Powell, S. R., Seethaler, P. M., Hamlett, C. L., Fuchs, D. (2010). The effects of schema-broadening instruction on second graders' word-problem performance and their ability to represent word problems with algebraic equations: A randomized control study. *Elementary School Journal, 110,* 446-463. doi:10.1086/651191

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: A five year prospective study. *Journal of Educational Psychology, 104,* 206–223. doi:10.1037/a0025398

Hamlett, C. L., Schumacher, R.F., & Fuchs, L.S. (2011). *2011 Number Line* adapted from Siegler et al. (2011). Available from L.S. Fuchs, 228 Peabody, Vanderbilt University, Nashville, TN 37203.

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

Hayes, A. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process monitoring [White paper]. http://www.personal.psu.edu/jxb14/M554/articles/process2012.pdf.

Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods, 41*, 924-936. doi:10.3758/BRM.41.3.924

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

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second grade arithmetic. *American Educational Research Journal, 30*, 393-425. doi: 10.3102/00028312030002393

Jitendra, A. K., & Star, J. A. (2012). An exploratory study contrasting high- and low-achieving students' percent word problem solving. *Learning and Individual Differences*, *22*, 151-158. doi:10.1016/j.lindif.2011.11.003

Jitendra, A. K., Star, J. A., Rodriguez, M., Lindell, M., & Someki, F. (2011). Improving students' proportional thinking using schema-based instruction. *Learning and Instruction*, *21*, 731-745. doi:10.1016/j.learninstruc.2011.04.002

Jitendra, A. K., Star, J. A., Starosta, K., Leh, J. M., Sood, S., Caskie, G., Hughes, C. L., & Mack, T. R. (2009). Improving seventh-grade students' learning of ratio and proportion: The role of schema-based instruction. *Contemporary Educational Psychology, 34,* 250-264. doi:10.1016/j.cedpsych.2009.06.001

Jordan, N. C., Hansen, N., Fuchs, L. S., Siegler, R. S., Gersten, R., & Micklos, D. (2013). Developmental predictors of fraction concepts and procedures. *Journal of Experimental Child Psychology, 116*, 45-58. doi:10.1016/j.jecp.2013.02.001

Kilpatrick, J., Swafford, J., & Findell, B., (Eds.) (2001). Adding it up: Helping children learn mathematics. Washington, DC: National Academy Press.

Kwon, K., Kumalasari, C.D., & Howland, J.L. (2011). Self-explanation prompts on problem-solving performance in an interactive learning environment. *Journal of Interactive Online Learning, 2*, 96-112. www.ncolr.org/jiol

Mazzocco, M. M. M., & Delvin, K. T. (2008). Parts and 'holes': Gaps in rational number sense among children with vs. children without mathematical learning disabilities. *Developmental Science, 11*, 681-691. doi: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. doi:10.1037/0012-1649.35.1.164

Montague, M. (2007). Self-regulation and mathematics instruction. *Learning Disability Research & Practice*, *22,* 75–83. doi: 10.1111/j.1540-5826.2007.00232.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2013). *Common Core State Standards (Mathematics Standards).* Washington DC: Author. http://www.corestandards.org/

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

Ni, Y. J. (2001). Semantic domains of rational number and acquisition of fraction equivalence. *Contemporary Educational Psychology, 26,* 400-417. doi:10.1006/ceps.2000.1072

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

Pickering, S., & Gathercole, S. (2001). *Working Memory Test Battery for Children.* London: The Psychological Corporation.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879-891. doi: 10.3758/BRM.40.3.879.

Resendez, M., & Azin, M. (2009). *A study on the effects of Pearson's 2009 enVisionMATH program. 2007–2008: First year report.* Jackson, WY: PRES Associates.

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development*, *77*, 1-15. doi: 10.1111/j.1467-8624.2006.00852.x

Rittle-Johnson, B., Fyfe, E.R., Loehr, A.M., & Miller, M.R. (2015). Beyond numeracy in preschool: Adding patterns to the equation. *Early Childhood Research Quarterly, 31*, 101-112. http://dx.doi.org/10.1016/j.ecresq.2015.01.005

Rittle-Johnson, B., Siegler, R. S., & Alibali, M.W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93,* 346-362. doi:10.1037/0022-0663.93.2.346

Schumacher, R. F., & Fuchs, L. S. (2012). Does understanding relational terminology mediate effects of intervention on compare word problems? *Journal of Experimental Psychology*, *111*, 607-628. doi:10.1016/j.jecp.2011.12.001

Schumacher, R. F., Namkung, J. M., Malone, A., & Fuchs, L. S. (2013). *2013 Fraction Battery-revised*. Available from L.S. Fuchs, 228 Peabody, Vanderbilt University, Nashville, TN 37203.

Schworm, S., & Renkl, A. (2002). Learning by solved example problems: Instructional explanations reduce self-explanation activity. In *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 816-821). Mahwah, NJ: Erlbaum.

Scott-Foresman Addision-Wesley (2011). enVisionMATH. San Antonio: Pearson.

Seaman, M. A., Levin, J. R., & Serlin, R. C. (l991). New developments in pairwise multiple comparisons: Some powerful and practical problems. *Psychological Bulletin, 110*, 577-586. doi:10.1037/0033-2909.110.3.577

Seethaler, P. M., Fuchs, L. S., Star, J. R., & Bryant, J. (2011). The cognitive predictors of computational skill with whole versus rational numbers: An exploratory study. *Learning and Individual Differences*, *21*, 536-542. doi:10.1016/j.lindif.2011.05.002

Siegler, R.S. (2002). Microgenetic studies of self-explanation. In N. Garnott & J. Parziale (Eds.), *Microdevelopment: A process-oriented perspective for studying development and learning* (pp. 31-58). Cambridge: Cambridge University Press.

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

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

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games – but not circular ones – improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101*, 545−560. doi:10.1037/a0014239

Swanson, H.L. (2014). Does cognitive strategy training on word problems compensate for working capacity in children with math difficulties? *Journal of Educational Psychology, 106*, 831-848. doi:10.1037/a0035838

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013). *Elementary school mathematics intervention report: enVisionMATH*. (http://ies.ed.gov/ncee/wwc/pdf/intervention_reports/wwc_envisionmath_011513.pdf

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.

Whitenack, J. W., & Yackel, E. (2002). Making mathematical arguments in the primary grades: The importance of explaining and justifying one's ideas. *Teaching Children Mathematics, 8*, 524-527.

Wilkinson, G. S., & Robertson, G. J. (2006). *Wide Range Achievement Test 4 professional manual*. Lutz, FL: Psychological Assessment Resources.

Wittrock, M.C. (1990). Generative processes of comprehension. *Educational Psychologist*, *24*, 345-376. doi:10.1207/s15326985ep2404_2

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities.* Itasca, IL: Riverside Publishing.

Table 1
*Curricular Emphases in Study Conditions* (*n* =52 teachers)

| Emphasis | Fraction Faceoff! EXP | Fraction Faceoff! WP | Control Group M (*SD*) |
|---|---|---|---|
| % Types of Fraction Representations Used | | | |
| Fraction tiles | 10 | 10 | 15.87 (15.46) |
| Fraction circles | 15 | 15 | 9.94 (10.33) |
| Pictorial representations with shaded regions | 5 | 5 | 33.94 (15.67) |
| Fraction blocks | 0 | 0 | 16.21 (13.06) |
| Number lines | 70 | 70 | 20.87 (11.06) |
| Other | 0 | 0 | 3.17 ( 7.67) |
| % Activities/Strategies for Comparing Fraction Magnitudes | | | |
| Cross multiplying | 0 | 0 | 21.15 (23.00) |
| Finding common denominator | 5 | 5 | 20.67 (16.42) |
| Thinking about relative placement on number lines | 30 | 30 | 12.40 ( 8.94) |
| Comparing fractions to benchmark fraction | 30 | 30 | 13.08 (11.38) |
| Drawing a picture of each fraction | 1 | 0 | 15.38 ( 9.54) |
| Using manipulatives | 5 | 5 | 7.12 ( 8.24) |
| Considering meaning of numerator &denominator | 29 | 30 | 8.17 ( 9.08) |
| Other | 0 | 0 | 2.03 ( 5.96) |
| % Explanation Strategies | | | |
| State explanations for answers | 5 | 0 | 22.06 (12.40) |
| Write important information | 3 | 0 | 19.65 (12.11) |
| Draw picture | 85 | 85 | 21.38 ( 9.48) |
| Label picture | 7 | 10 | 14.17 ( 8.90) |
| Make a table | 0 | 0 | 6.95 ( 7.24) |
| Write an equation | 0 | 0 | 15.79 ( 8.01) |
| Other | 0 | 0 | 0.00 ( 0.00) |
| % Explanations | | | |
| Oral | 80 | 100 | 46.44 (17.47) |
| Written | 20 | 0 | 53.56 (17.47) |
| % Word-Problem Strategies | | | |
| Draw a picture | NA | 4 | 24.27 (13.35) |
| Make a table | NA | 0 | 7.64 ( 8.17) |
| Make an array | NA | 32 | 0.00 ( 0.00) |
| Write a number sentence | NA | 0 | 18.77 ( 8.84) |
| Use words to explain thinking | NA | 32 | 19.27( 6.65) |
| Relies on key words | NA | 0 | 20.13(12.72) |
| Identifies problem within a problem type | NA | 32 | 9.92( 9.70) |
| Other | NA | 0 | 0.00( 0.00) |

Table 2

*Pretest and Posttest Fraction Scores by Condition, with Not-At-Risk Norms and Achievement Gaps*

| | At-Risk Condition | | | | | | | | Achievement Gap (Effect Size[a]) Not-at-Risk vs. | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | WP (n=69) | | Explain (n=73) | | Control (n=70) | | Not-at-Risk (n=320) | | | | |
| Variable | Mean | *(SD/SE[b])* | Mean | *(SD/SE)* | Mean | *(SD/SE)* | Mean | *(SD)* | WP | Explain | Control |
| Magnitude Comparison/Explanation Measure | | | | | | | | | | | |
| Correct Comparisons-Post | 5.94 | (1.38) | 6.47 | (1.09) | 4.59 | (1.65) | -- | | | | |
| Explanation Quality-Post | 2.93 | (3.89) | 9.10 | (9.30) | 1.14 | (2.09) | -- | | -- | -- | -- |
| Word Problems | | | | | | | | | | | |
| Pre | 4.83 | (4.67) | 4.66 | (3.61) | 4.74 | (4.06) | 9.48 | (5.38) | -0.86 | -0.89 | -0.88 |
| Post | 13.61 | (6.48) | 5.88 | (3.84) | 6.49 | (5.37) | 13.52 | (6.46) | +0.01 | -1.18 | -1.09 |
| Adjusted-Post | 13.57 | (0.59) | 5.91 | (0.58) | 6.48 | (0.59) | -- | | | | |
| NAEP | | | | | | | | | | | |
| Pre | 9.71 | (3.55) | 10.14 | (2.78) | 10.50 | (3.33) | 14.48 | (4.18) | -1.14 | -1.03 | -0.95 |
| Post | 14.88 | (4.20) | 14.45 | (3.35) | 12.68 | (3.85) | 18.95 | (3.81) | -1.07 | -1.18 | -1.64 |
| Adjusted-Post | 15.19 | (0.36) | 14.44 | (0.35) | 12.39 | (0.35) | -- | | | | |
| Number Line | | | | | | | | | | | |
| Pre | 0.29 | (0.08) | 0.31 | (0.07) | 0.31 | (0.09) | -- | | | | |
| Post | 0.20 | (0.09) | 0.21 | (0.08) | 0.26 | (0.08) | -- | | -- | -- | -- |
| Adjusted-Post | 0.20 | (0.01) | 0.21 | (0.01) | 0.26 | (0.01) | -- | | | | |
| Calculations | | | | | | | | | | | |
| Pre | 5.57 | (4.35) | 5.05 | (4.59) | 4.74 | (4.58) | 8.35 | (4.85) | -0.57 | -0.68 | -0.74 |
| Post | 21.75 | (7.10) | 20.42 | (6.63) | 8.89 | (4.92) | 19.21 | (9.25) | +0.27 | +0.13 | -1.12 |
| Adjusted-Post | 21.59 | (0.73) | 20.45 | (0.71) | 9.02 | (0.73) | -- | | | | |

[a]Effect size achievement gaps are difference in posttest scores, divided by not-at-risk *SDs*.

[b]*SE* is standard error, reported for adjusted posttest scores; *SD* is standard deviation.

*Fraction Number Line* (0-2) is Schumacher et al. (2011). *NAEP* is National Assessment of Educational Progress items (19 easy, medium, and hard fourth-grade and easy eighth-grade released fraction items). Correct Comparisons and Explanation Quality are from *Magnitude Comparisons/Explanation Quality* in the Fraction Battery-revised (Schumacher et al., 2013). Word Problems is *Multiplicative Word Problems* from the Fraction Battery-revised (Schumacher et al., 2013).

Table 3

*Intervention Effects, Follow-Up Tests, and Effect Sizes*

| | Outcome | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mag Compare Correct | | Explain Quality | | WPs | | NAEP | | Number Line | | Calculations | |
| Corrected Model | 34.66 | (<.001) | 36.10 | (<.001) | 28.88 | (<.001) | 53.30 | (<.001) | 16.68 | (<.001) | 67.97 | (<.001) |
| Intercept | 3526.47 | (<.001) | 116.44 | (<.001) | 144.51 | (<.001) | 92.74 | (<.001) | 33.48 | (<.001) | 572.02 | (<.001) |
| Pretest | NA | | NA | | 35.91 | (<.001) | 138.06 | (<.001) | 24.59 | (<.001) | 14.53 | (<.001) |
| Condition | 34.66 | (<.001) | 16.52 | (.005) | 23.93 | (<.001) | 16.52 | (.005) | 10.70 | (<.001) | 91.07 | (<.001) |
| Pretest x Condition | NA | | NA | | 3.77 | (.155) | NA | | NA | | NA | |
| Follow-Up[b] | 2>1>3 | | 2>1>3 | | 1>2=3 | | 1>2=3 | | 1>2=3 | | 2=1>3 | |
| Effect Sizes[c] | | | | | | | | | | | | |
| Explain v. Control | 1.37 | | 1.37 | | -0.12 | | 0.57 | | 0.63 | | 1.98 | |
| WP v. Control | 0.89 | | 0.60 | | 1.20 | | 0.70 | | 0.71 | | 2.08 | |
| WP v. Explain | -0.43 | | -0.93 | | 1.48 | | 0.20 | | 0.12 | | 0.16 | |

[a]For Number Line, NAEP, and Calculations, $F(3,208)$ for corrected model; $F(1,208)$ for intercept; $F(1,208)$ for pretest; and $F(2,208)$ for condition. For Word Problems, $F(5,206)$ for corrected model; $F(1,206)$ for intercept; $F(1,206)$ for pretest; $F(2,206)$ for condition; and $F(2,206)$ for pretest x condition. For Magnitude Comparison-Correct and Explain-Quality, $F(2,209)$ for corrected model; $F(1,209)$ for intercept; and $F(2,209)$ for condition.

[b]1=WP; 2=Explain; 3=Control.

[c]Positive effect sizes indicate stronger performance (the Number Line effect size were multiplied by -1). ES is the difference between adjusted posttest means divided by the pooled *SD* of the unadjusted posttest scores. For the Explain variables, ES is the difference between posttest means divided by the pooled *SD* of the posttest scores

*Fraction Number Line* (0-2) is Schumacher et al. (2011). *NAEP* is National Assessment of Educational Progress items (19 easy, medium, and hard fourth-grade and easy eighth-grade released fraction items). Mag Compare and Explain Quality are number of correct comparisons and explanation quality from the *Magnitude Comparison/Explanation Quality* measure in the *Fraction Battery*-revised (Schumacher et al., 2013). WPs is *Multiplicative Word Problems* from the *Fraction Battery*-revised (Schumacher et al., 2013).

Table 4

*Descriptive Information and Correlations for Moderation Analysis Variables for Explanation and Word-Problem Condition Students (n=142)*

| | Raw Scores | | Standard Scores | | Correlations | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Variables | Mean | *(SD)* | Mean | *(SD)* | W | R | L | M | WP1 |
| Working Memory (W) | 16.65 | (4.24) | 77.25 | (14.96) | | | | | |
| Reasoning (R) | 15.99 | (6.16) | 44.75 | (10.18) | **.26** | | | | |
| Language Comprehension (L) | 20.92 | (3.39) | 89.57 | (13.40) | **.37** | **.18** | | | |
| Magnitude Compare– Post (M) | 6.20 | (1.28) | NA | | .14 | .08 | .07 | | |
| Word Problems – Pre (WP1) | 4.74 | (4.14) | NA | | .16 | **.25** | **.23** | .12 | |
| – Post | 9.63 | (6.54) | NA | | **.21** | .15 | **.26** | .02 | **.29** |

Bolded correlations are significant ($p < .05$). Working memory is *Working Memory Test Battery for Children - Counting Recall* (Pickering & Gathercole, 2001). Reasoning is *WASI Matrix Reasoning* (Wechsler, 1999). Language comprehension is *Woodcock Diagnostic Reading Battery - Listening Comprehension* (Woodcock et al., 2001). Magnitude Compare is from the *Magnitude Comparison/Explanation Quality* measure from the Fraction Battery-revised (Schumacher et al., 2013). Word problems is *Multiplicative Word Problems* from the Fraction Battery-revised (Schumacher et al., 2013).

Note. Correlations between some moderator variables (which were measured at the start of the study) and outcomes (which were measured at the end of the study) are low because, over the course of the study, intervention was delivered to disturb the natural relation between these variables, but only for half the sample - for the condition that received relevant intervention. Across the two conditions (as reported in this table), this reduces the magnitude of these relations.

Table 5
*Moderation Analyses*

| | Outcome/Moderator | | | | | |
|---|---|---|---|---|---|---|
| | Magnitude Comparison | | | Word Problems | | |
| | Working Memory | Language Comprehension | Reasoning | Working Memory | Language Comprehension | Reasoning |
| Effects | *t (p)* | *t (p)* | *t (p)* | *t (p)* | *t (p)* | *t (p)* |
| Constant | -0.11 (.914) | 0.00 (.998) | 0.02 (.981) | -0.06 (.951) | 0.19 (.849) | -0.56 (.575) |
| Moderator | -2.03 (.044) | 0.17 (.865) | 1.11 (.268) | 1.01 (.315) | 2.23 (.027) | 0.52 (.605) |
| Condition | 4.97 (<.001) | 4.93 (<.008) | 4.94 (<.001) | 6.08 (<.001) | 5.74 (<.001) | 6.57 (<.001) |
| Moderator x Condition | -2.02 (.047) | 0.15 (.879) | -0.47 (.640) | -1.62 (.108) | 0.00 (.996) | -3.07 (.003) |
| Control Variables | | | | | | |
| Pretest WPs | NA | NA | NA | 1.21 (.228) | 1.35 (.181) | 0.59 (.553) |
| Pre-WPs x Condition | NA | NA | NA | 0.04 (.971) | -0.20 (.843) | 0.62 (.536) |
| Working Memory | NA | -1.71 (.089) | -1.73 (.086) | NA | 1.30 (.182) | 1.20 (.233) |
| Language Comp | 0.07 (.947) | NA | 0.22 (.823) | 2.13 (.035) | NA | 2.62 (.010) |
| Reasoning | 1.18 (.239) | 1.25 (.213) | NA | 1.10 (.273) | 1.17 (.243) | NA |

Working memory is *Working Memory Test Battery for Children - Counting Recall* (Pickering & Gathercole, 2001). Language comprehension is *Woodcock Diagnostic Reading Battery - Listening Comprehension* (Woodcock et al., 2001). Reasoning is *WASI Matrix Reasoning* (Wechsler, 1999). Magnitude Comparison is from the *Magnitude Comparison/Explanation Quality* measure from the Fraction Battery-revised (Schumacher et al., 2013). Word problems is *Multiplicative Word Problems* from the Fraction Battery-revised (Schumacher et al., 2013).

Figure 1. *Timeline of intervention focus in EXP, WP, and Core (the multi-component program) by study week.*

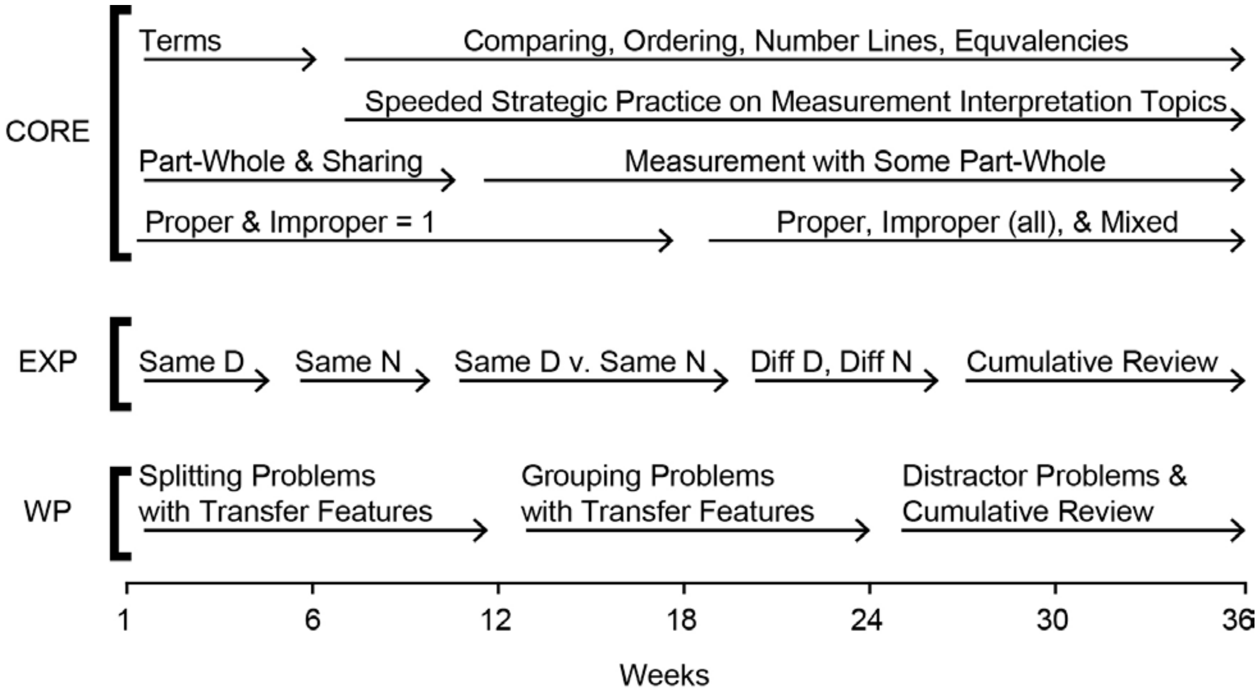

Figure 2

*Visualization of the statistical interaction showing how students' incoming working memory capacity (x-axis) moderates the effect of intervention condition (EXP scores in white bars versus WP scores in black bars) on the explanation quality outcome.*

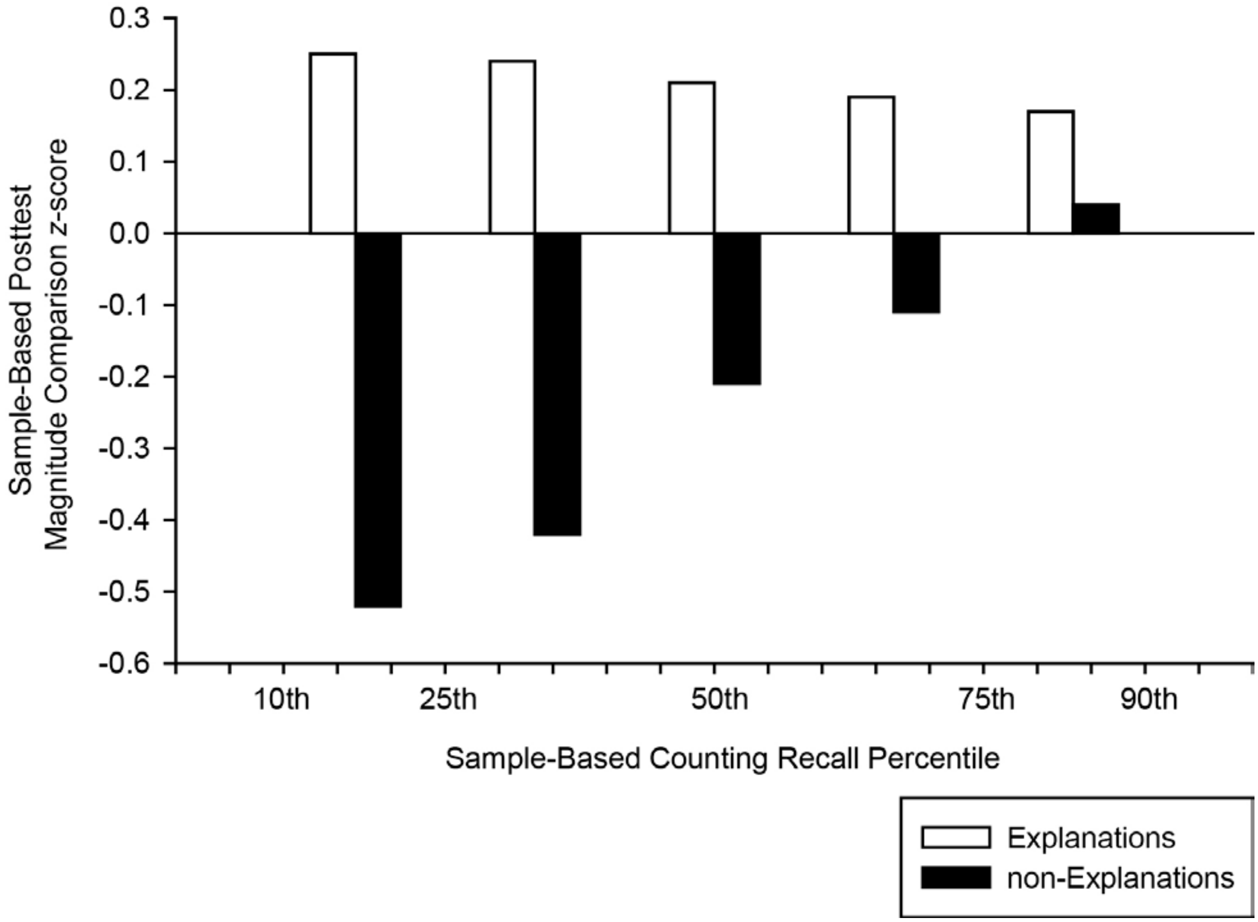

Figure 3

*Visualization of the statistical interaction showing how students' incoming reasoning ability (x-axis) moderates the effect of intervention condition (WP scores in white bars versus EXP scores in black bars) on the word-problem outcome.*