

Can Argumentative Writing Improve Math Knowledge for Elementary Students with a Mathematics Learning Disability?: A Single-Case Classroom Intervention Investigation



Sharlene A. Kiuahara, Joel R. Levin, Malynda Tolbert, Megan Erickson, and Kenny Kruse

Abstract Students with high incidence disabilities continue to perform considerably lower than their same-aged peers without disabilities in the areas of written expression and mathematical reasoning. This is especially concerning for students who come from diverse cultural, linguistic, and socio-economic backgrounds. We examined the effectiveness of a writing-to-learn mathematics intervention designed for students with a mathematics disability. The intervention incorporated the six-stages of Self-Regulated Strategy Development (SRSD) that targeted students' understanding of fractions as numbers and their argumentative writing and mathematical reasoning. A single-case multiple-baseline design was implemented with seven special education teachers who were randomly assigned to the staggered tiers of the design. Following 2 days of professional development and training, the teachers initiated the intervention in their classrooms. Visual and statistical analyses of the data revealed selected positive baseline-to-intervention phase changes in students' performance during implementation of SRSD. Implications and future directions of the research are discussed.

Keywords Argument writing · Writing-to-learn mathematics · Self-regulation strategy development

S. A. Kiuahara (✉) · M. Tolbert · M. Erickson · K. Kruse
University of Utah, Salt Lake City, UT, USA
e-mail: s.kiuahara@utah.edu

J. R. Levin
University of Arizona, Tucson, AZ, USA

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy
Studies 25, https://doi.org/10.1007/978-3-031-36472-3_11

1 Introduction

A recent meta-analysis conducted by Graham et al. (2020) has demonstrated that writing as a learning activity increases content-area learning for students in Grades 1–12, including mathematics. When learning mathematics, Resnick (1987) re-characterized the “3-Rs” as reasoning, writing, and arithmetic, as essential skills for children to carry out a series of steps for solving a mathematical problem flexibly and accurately. The National Governors Association and the Council of Chief State School Officers also recognized the importance of developing students’ reasoning and language when learning mathematics and included eight practices in the Common Core State Standards for Mathematics (CCSS-M) to develop students’ mathematical expertise (CCSS-M, 2010). Such activities involve constructing arguments, communicating their reasoning using clear definitions and explanations to justify their answers, and critiquing their peers’ reasoning. Thus, students use language to become active agents in constructing new knowledge (Boscolo & Mason, 2001; Newell, 2006) and are likely to show larger gains in their problem-solving performance (Ball & Bass, 2003; National Mathematics Advisory Panel [NMAP], 2008).

In this chapter, we report a mathematics-intervention study that extends two feasibility studies presented by Hacker et al. (2019) and Kiuvara et al. (2020) in which the planning and constructing of written arguments facilitated the fraction learning and quality of mathematical reasoning of 5th and 6th-grade students with a mathematics learning disability (MLD). Although their previous research targeted students with and at-risk for a mathematics learning disability (MLD), the students in the current study were classified as having a MLD and were receiving specialized instruction in mathematics.

1.1 *What Are Barriers for Students with MLD When Learning Fractions?*

Developing foundational knowledge about fractions is an essential building block for developing algebraic reasoning and predicting success for learning secondary mathematics (Bailey et al., 2015; Siegler et al., 2012). Proficiency with fractions requires students to make a conceptual shift from understanding whole numbers to two quantities that convey a single numerical value (e.g., $\frac{2}{3}$). This shift views the magnitude of the numerator and denominator as a unit rather than as separate numbers (Fuchs et al., 2013; Jordan et al., 2013). Although many 4th-graders fail to perform at or above the 25th percentile in mathematics, students with disabilities continue to score *below basic proficiency levels* compared to their same-aged peers without disabilities (National Center for Educational Statistics, 2015). Research has shown that students with MLD also experience difficulty using mathematical notation and accurately solving fraction problems involving multiple steps (Bryant &

Bryant, 2008; Geary, 2011). Students with MLD continue to struggle learning fractions through middle school (Mazzocco et al., 2013), which places them even further behind their same-grade peers without disabilities.

1.2 A Place for Writing Instruction in Math Class?

Researchers have found that writing about one's learning promotes deeper engagement and active reasoning about new ideas (Bangert-Drowns et al., 2004; Graham et al., 2020; Hubner et al., 2006). Children use language to make sense and meaning of mathematical content, and language ability, in turn, supports and reinforces the conceptual knowledge needed for learning mathematics (Desoete, 2015; Vukovic & Lesaux, 2013). However, research has also indicated that students with MLD (a) may exhibit comorbidity with language development and learning mathematics (Korhonen et al., 2012); (b) often struggle with reasoning and communicating their ideas (NAEP, 2011); (c) often have difficulty with working memory, processing speed, and self-regulation when approaching and completing a task (Geary, 2011; Jitendra & Star, 2011; Jordan et al., 2013); and (d) have limited background and vocabulary knowledge for explaining or justifying their solutions and the solutions of their peers (Gersten et al., 2009; Krowka & Fuchs, 2017; Woodward et al., 2012). For students who may have MLD and speak English as a second language, using the process of writing to develop targeted academic vocabulary is essential (Cuenca-Carlino et al., 2018).

Writing provides students with a permanent record of their thinking (Hacker & Dunlosky, 2003). Therefore, writing about math learning may be helpful to students with MLD because writing allows students to reformulate and make sense of complex mathematical concepts. Engaging students in activities for constructing logical arguments and evaluating the logic or reasoning of peers are considered mathematical practices that develop mathematical expertise (CCSS-M, 2010; NMAP, 2008). However, some may argue that combining writing and mathematics may create further barriers to learning mathematics, especially for children with MLD.

1.3 Writing Strategies Instruction and Writing-to-Learn Mathematics

Because we are interested in the benefits of using argumentative writing as a learning tool to develop mathematical knowledge and reasoning for students with MLD, we drew upon the extensive Self-Regulated Strategy Development (SRSD) evidence base that sequences explicit writing and self-regulation strategies for students with and without mild to moderate support needs (see Graham et al., 2012; Gillespie & Graham, 2014; Hebert & Powell, 2016; Harris & Graham, 2009). Briefly

described here, SRSD consists of six stages of instruction (Develop Background Knowledge, Discuss It, Model It, Memorize It, Support It, and Independent Performance) to help students independently manage their learning and writing process. The six stages focus on developing students' background knowledge for understanding the purpose, value, and characteristics of the strategies. Students learn to self-regulate and self-monitor their learning process, memorize the learning strategies, and give and receive feedback on their writing. Together with the teacher or other students, students actively engage in discussions, learn the terminology and vocabulary needed to articulate their mathematical understandings, and carry out the rhetorical structures of a specific writing genre, such as argumentation. During instruction, the stages are taught recursively, so the teacher can repeat stages and differentiate instruction to meet individual students' learning needs. The teacher models how to use the learning and self-regulation strategies and facilitates students' learning until each student can use the strategies independently and support their peers in their learning and writing process (Kiuahara et al., 2020).

Two strategies called FACT + R²C² (herein referred to here as FACT) were designed to help students with and at-risk for MLD reason through a mathematics problem while constructing a written argument using the six stages of SRSD. FACT represents the following steps: F = *Figure* out plan (What is my task? Do I understand the problem? What do I need to know? What tools do I need?); A = *Act* on it (What reasons, evidence and support will I use? What words will I choose? How will I interpret my results?); C = *Compare* my reasoning with a peer's (What is similar or different? What are my reasons? Does it make sense? Can we make improvements?); and T = *Tie* it up in an argument which prompts students to go through the steps of RRCC. R²C² represents the following steps: R = Did I *restate* the task?; R = Did I provide *reasons*, evidence, and support?; C = Did I provide a *counterclaim* that addresses an answer different from my own?; and C = Did I wrap it up with a *concluding* statement?

FACT was empirically tested in studies by Hacker et al. (2019) and Kiuahara et al. (2020). Kiuahara and colleagues initially tested the effects of FACT using a pre-posttest cluster-randomized controlled trial in which 10 teachers were randomly assigned to the FACT or business as usual conditions by teacher type (i.e., special educator or general educator) and grade (i.e., 4th, 5th, or 6th grades). The teachers in the FACT condition received 2 days of professional development (PD) before implementing the FACT lessons with their students. Treatment fidelity observations were conducted across 33% of the class sessions and was high (96% across all teachers, range 89–100%). The outcomes from pretest to posttest favored students in the FACT condition on a fraction test (Hedges' $g = 0.60$), quality of mathematical reasoning ($g = 1.82$); the number of argumentative elements ($g = 3.20$), and total words written ($g = 1.92$). We found that students with MLD in the FACT condition demonstrated greater gains in fraction scores from pretest to posttest than students without MLD ($n = 12$, $g = 1.04$). These findings showed promise for implementing a writing-to-learn math intervention in which students with MLD constructed arguments and critiqued the reasoning of their peers.

The study by Hacker et al. (2019) tested the effects of FACT on students with MLD using a single-case multiple baseline design (MBD) with an associated randomization test (Levin et al., 2018) to make informed decisions for further development of the lessons (Levin, 1992). Five special education teachers from different schools implemented the intervention with their 5th- and 6th-grade students ($n = 34$). The teachers received 2 days of PD before they were randomized to begin instruction at staggered points in time. A pre- posttest fraction measure indicated an average increase from the beginning to the end of the study ($d = 0.70$). However, Parker et al.'s (2014) rescaled Nonoverlap of All Pairs (NAP) effect-size indices (Gafurov & Levin, 2022) at the classroom-level produced mixed results for fraction accuracy, .32 (range, .13 to .69); mathematical reasoning, .61 (.08 to .97); number of rhetorical elements, .45 (.15 to .81); and total words written, .14 (.54 to .78). Teachers were observed every third lesson. Although overall treatment fidelity across the teachers was high (87%), some teachers were showing stronger student gains than others and some teachers showed higher treatment fidelity than others, ranging from 66% to 99%. Thus, further examination was needed to increase treatment fidelity and to identify the skills and knowledge needed for special education teachers to implement content-rich language and mathematics instruction.

1.4 Purpose of the Present Study

The purpose of the present study was to build from the efficacy studies presented earlier by Hacker et al. (2019) with a similar single-case randomized design MBD and associated statistical analysis. Doing so here provided us with data to identify areas for additional refinement to the FACT lessons and the Writing-to-Learn PD protocol. The primary research question that guided our study was: To what extent do students with MLD who receive the FACT intervention demonstrate gains in fraction knowledge? We also wanted to understand the effects of the intervention on their quality of mathematical reasoning, argumentative elements, and total words written. We predicted that using argument writing as a tool during learning would encourage students to be more precise in communicating their mathematical reasoning (Graham et al., 2020; Resnick, 1987).

2 Method

2.1 Setting and Participants

This study took place in a large and ethnically diverse school district located in the intermountain region of the United States. After receiving institutional review board approval to conduct this study, we contacted key district personnel who identified

licensed special education teachers. Eight special education teachers from eight different elementary schools consented to participate, and three sequential forms of teacher randomization took place prior to the start of the study. First, eight teachers were randomly assigned to four dyads. Then, each dyad was randomly assigned to predetermined dates during which the pair of teachers participated in 2 days of PD before implementing the FACT intervention. Finally, each teacher from the same dyad was randomly assigned to one of two adjacent dates within the eight intervention tiers of the MB design.

Absenteeism presented an unforeseen challenge for one teacher because of a family emergency, which considerably delayed the second of 2 days of her initial PD phase of the study (discussed below), along with her assigned intervention start date and continuing intervention-day absences. Therefore, that teacher and her students ($n = 4$) were not considered to be part of the formal study's data analysis and so the following results and discussion are based on the remaining seven teachers' classes.

Teachers provided specialized mathematics instruction in small groups to 5th and 6th-grade students with MLD for 45 min per day, four times per week. Five schools qualified for Title 1 services. The teachers had taught on average for 13.38 years (range = 5–27 years). Four teachers had a master's degree, two had a bachelor's degree, and one had a Juris doctorate. All teachers were female.

The 5th and 6th-grade students from each teacher's specialized mathematics class were invited to participate in the study if they (a) had at least one mathematics learning goal on their Individualized Education Plan (IEP), (b) were receiving specialized instruction in mathematics, and (c) were able to write a complete sentence on a standardized writing test.

The student participants ($n = 27$) included 15 (56%) 5th-graders, 12 (44%) 6th-graders, 14 (52%) male, and 10 (37%) English learners (ELs). Fifteen students (56%) were Latinx, 8 (30%) were White, 2 (7%) were Black, and 2 (7%) were multiracial. The students were administered two screening measures: (a) the mathematics subtest of the *Wide Range Achievement Test, 4th Ed. (WRAT-4)* (Wilkinson & Robertson, 2006) and (b) a writing subtest from *Wechsler Individual Achievement Test (WIAT), 3rd Ed.* (Wechsler, 2009), in which students wrote a short essay response to an expository prompt. Students' writing was scored for word count, theme development, and text organization. All students scored below the 14th percentile on the mathematics subtest ($M = 6.92$; range = 3.17–13.36). The average students' percentile ranking on the written expression subtest was 6.90 (range = 3.17–13.67).

2.2 Single-Case Intervention Design

A nonconcurrent MBD, with random assignment of classrooms to the staggered multiple-baseline levels (or "tiers") – a single-case design with a high degree of scientific credibility (Kratochwill & Levin, 2010; Levin, 1992; Levin et al., 2018; Slocum et al., 2022) – was adopted to evaluate the effectiveness of the intervention.

A MBD was implemented because the skills and knowledge that students may have acquired through participation were unlikely to be reversed, and this single-case intervention design allowed each teacher's class to serve as its own control (e.g., Horner & Odom, 2014; Kiuahara et al., 2017). The nonconcurrency of the design was necessitated by the irregularities of school-schedule timing, teacher absences, weather conditions, etc., for the commencement of each teacher's intervention on pre-specified calendar dates. Hence, the to-be-reported stagger of the design is represented by the number of pre- and post-intervention "sessions" rather than by actual chronological dates (Slocum et al., 2022). The purpose of implementing randomization in the design and associated statistical analysis was: (a) to improve the internal validity of the study, thereby providing a small-scale proxy to a randomized controlled trial (Kratochwill & Levin, 2010); and (b) to allow for a formal statistical assessment of the intervention's effectiveness based on a well-controlled Type 1 error probability while furnishing adequate statistical power for uncovering the intervention effects of interest (Levin et al., 2018).

The PD was provided by the third and fourth authors prior to the teachers' intervention start date. The fractions and argumentative writing components were guided by Hacker et al.'s (2019) and Kiuahara et al.'s (2020) PD protocols, which included PD components needed for teachers to "buy in" and effectively implement the intervention (e.g., multiple opportunities for active learning through modeling and practice and incorporating the same materials during PD that students have used previously) (Harris et al., 2012, 2014). We added discussion and activities that centered around developing and extending teachers' conceptual and procedural understandings of fractions, common misconceptions students exhibit when learning fractions, and using questioning to better engage students in learning mathematics (Borko et al., 2015; Polly et al., 2014; Jayanthi et al., 2017). The teachers attended the PD in pairs to encourage collective participation in a safe learning environment and allowed PD facilitators to provide feedback and establish rapport for lending support to teachers once the teachers began instruction during the intervention phase.

2.3 Intervention

The FACT intervention consisted of five lessons using the six stages of SRSD described earlier (i.e., Develop Background Knowledge, Discuss It, Memorize It, Model It, Support It, and Independence), which embedded explicit instruction for improving students' writing knowledge and performance, self-efficacy, and strategic behavior with their writing and learning processes (Harris & Graham, 2009). We situated the language content of the FACT lessons around four mathematical practices: (a) construct arguments and critique the reasoning of peers, (b) make sense of problems and persevere in solving them, (c) use appropriate tools strategically, and (d) attend to precision by speaking and writing with precise mathematics vocabulary, describing relationships clearly, and calculating problems accurately (CCSS-M, 2010). The fraction content included understanding equivalence, comparing

fractions that refer to the same whole, composing and decomposing fractions, adding and subtracting with like and unlike denominators, and using equivalent fractions to solve problems with unlike denominators. The writing content included writing arguments to support claims with reasons and evidence. The fraction content was taught using a sequence of multiple representations (e.g., fraction blocks, number lines, area models, and numerical and mathematical notation) (Hughes et al., 2014; Witzel et al., 2003).

2.4 Measures

Distal Fraction Measure

The *easyCBM* Math Number and Operations assessment (Tindal & Alonzo, 2012) was modified to include 27 fraction items (0–27 points possible). The test items consisted of multiple-choice questions that focused on magnitude, equivalence, comparing two fractions and adding and subtracting fractions with like and unlike denominators. Two equivalent forms of the test were counterbalanced and administered to students before and at the end of the intervention phase and at the end of the intervention and post-intervention. Two independent scorers scored 100% on all assessments. Interrater reliability (IRR) between two scorers was 100%.

Progress Monitoring Fractions

Twenty equivalent fraction probes were developed and administered to students each week during the study. Each fraction probe consisted of 14 questions (1 point each) and was divided into three sections: (a) placing fractions on a number line; (b) comparing the magnitude of two fractions; and (c) computational accuracy for adding and subtracting two fractions. Two independent scorers scored 100% on all assessments with an IRR of 100%.

Progress Monitoring Writing

Twenty equivalent writing probes were administered weekly to measure students' ability to construct an argumentative paragraph in which students justified their solution to a fraction problem during untimed conditions. The students' written responses were scored for quality of mathematical reasoning, argumentative writing elements, and total words written following the procedures used by Hacker et al. (2019) and Kiuvara et al. (2020). Two independent scorers scored 100% of all assessments.

Quality of Mathematical Reasoning

Students' papers were scored holistically for the quality of mathematical reasoning following the procedures outlined in Kiuahara et al. (2020). The scoring index was from 0 to 12, with higher scores indicating higher reasoning quality and computational accuracy. For example, a score of 0 indicated that the student wrote no response or showed no understanding of the problem (e.g., I don't know). A score of 5 or 6 indicated a student solved the problem correctly but had gaps in reasoning to support their answer or provided little support in a counterclaim. The responses that scored 11 or 12 included a clear and focused understanding of the problem, an accurate and fully supported position, a supported counterclaim, and controlled writing with sequencing and strong transitions. Two independent scorers scored 100% of all assessments. Disagreements of more than + or -3 points between the two scorers were resolved by discussion. IRR was 90%.

Argumentative Elements

Students' papers were scored for six argumentative elements (0–36 points) following the procedures from Kiuahara et al. (2020). The papers were scored for the following: (a) included a statement that represented the mathematics task (e.g., My task is to compare the fractions $\frac{1}{4}$ and $\frac{2}{3}$.); (b) stated a claim or answer to the mathematics problem (e.g., I think that $\frac{2}{3}$ is greater than $\frac{1}{4}$.); (c) provided reasons and elaborations to support the claim (e.g., I used a number line, and $\frac{2}{3}$ is closer to one whole.); (d) provided a counterclaim or an incorrect solution to the problem (e.g., Others may argue that $\frac{1}{4}$ is greater $\frac{2}{3}$.); (e) provided reasons and elaborations to support the counterclaim (e.g., My peer might think $\frac{1}{4}$ is greater because 4 is greater than 3); and (f) provided a concluding statement (e.g., However, the number line shows that $\frac{2}{3}$ is greater than $\frac{1}{4}$ because it is closer the 1). Two independent scorers scored 100% of all assessments. Disagreements of more than + or -3 points between the two scorers were resolved by discussion. IRR was 96%.

Total Words Written

Students' writing was scored for total words written following the procedures outlined by Kiuahara et al. (2020). The third author typed verbatim the students' writing probes into a word processing program to eliminate bias for handwriting, spelling, and grammar errors (Kiuahara et al., 2012). A second scorer checked the typed probes for accuracy and resolved any differences with the first scorer. The word-processing program calculated the total number of words written.

2.5 *Treatment Fidelity*

Each teacher was observed for a minimum of 33% of instructional sessions during the intervention phase of the study. Based on our findings from the MBD study reported by Hacker et al. (2019), we established that a teacher required further coaching and support if their treatment fidelity during the observed session did not reach a 90% criterion. Across all teachers, the average percent of instructional components that were in agreement between two independent raters was 92% with a range of 83–97%. On average, the teachers completed the intervention in 28 instructional days (range = 16–42 days).

2.6 *Approach to Analysis*

As was already noted, a randomized single-case MBD and associated randomization tests (Levin et al., 2018) were adopted for the present study. Following those procedures, the eight participating teachers were randomly assigned to the design's staggered tiers, with a planned two outcome-observation stagger between tiers. The resulting design was implemented over a 16-week period and assured a minimum of 5 baseline (A-phase) observations and 4 intervention (B-phase) outcome observations for each teacher/classroom (but see the first paragraph of the immediately following section). All randomization tests were based on the average outcome performance of each teacher's class, which ranged in size from two to seven students, and those tests were directional (*viz.*, positing that the mean of the intervention phase would exceed the mean of the baseline phase) based on a Type I error probability of .05. In addition, because previous related research (e.g., Kiuvara et al., 2020) suggested that any expected intervention effects would not emerge in an immediate fashion, a two-outcome observation delay was built into the randomization-test analyses of all measures of mean between-phase change, based on the "data-shifting" procedure of Levin et al. (2017, p. 24) as operationalized in Gafurov and Levin's (2022) freely available *ExpRT* single-case randomization-test package.

3 **Summary of Results**

The to-be-summarized randomization-test analyses were conducted on five weekly outcome measures: Fractions Test, Mathematical Reasoning, Argumentative Elements, and Total Words Written. In addition, more fine-grained analyses were conducted on just the intervention's Lesson 1 and Lesson 2 content, which aligned with the first three stages of SRSD instruction (i.e., Develop Background Knowledge, Discuss It, Memorize It). As an overview, the analyses yielded mixed results, which in turn lead to inconclusive interpretations. This is attributable in part to the present

MBD's actual between-tier stagger of only one outcome assessment per week (equivalent to four instructional lessons, depending on the teaching pace of the teacher) rather than the planned-for two staggers (eight instructional lessons). That was an unfortunate consequence of the aforementioned intervention-scheduling constraints for the 7 participating teachers. As a result, less clear differentiation among the present study's staggered tiers serves to reduce the typically high scientific credibility of a randomized MBD (Levin et al., 2018, Figs. 2 and 3, pp. 298–300). An illustrative graph depicting the results for one of the outcome measures, Mathematical Reasoning, is presented in Fig. 1, with the two-observation delayed phases labeled Phase A* and Phase B*.

Primary analyses. Wampold and Worsham's (1986) MBD randomization-test procedure, the most appropriate and statistically powerful MBD randomization test available (Levin et al., 2018), was applied to the four primary outcome measures but did not produce any statistically significant A- to B-phase improvements (p -values ranging from .94 to .16). That said, there was a large *general* between-phase statistical performance increase on three of the four measures, according to Busk and Serlin's (1992) "no-assumptions" average d s and Parker et al.'s (2014) average NAPs. After rescaling the NAPs to range from 0.0 to 1.00, they represent the average (across-classrooms) proportion of observations in the seven classrooms' baseline and intervention phases that do not overlap (for Mathematical Reasoning: $d = 6.86$, NAP = .74; for Argumentative Elements; $d = 6.14$, NAP = .78; and for Total Words Written: $d = 2.45$, NAP = .55). At the same time, a "visual analysis" (Kratochwill et al., 2021) of the outcome data revealed that four of the seven classes' observed improvements were not synchronous with the introduction of the intervention (see the Mathematical Reading outcomes for Teachers 1, 3, 4, and 6 in Fig. 1). As such, those improvements could have been attributable, at least to some extent, to the students' year-long growth spurred by teachers' provision of regular mathematics instruction during the baseline phase, which varied across the staggered intervention start dates. It is worth noting that the apparent precipitous decline and recovery of the intervention-phase mean for Teacher 5's class is based on the performance of only two students.

Two other sets of statistical analyses shed additional light on the efficacy of the FACT intervention. In one, analyses were conducted to examine whether the just-noted findings were consistent with another aspect of MBD logic, namely, that as each tier in the design exhibits an A- to B-phase improvement, the lower-level tiers do not – consistent with what has been referred to as a "vertical analysis" (see Kratochwill et al., 2021). An illustration of a differentiated vertical stair-step pattern can be appreciated through an examination of the Mathematical Reading measure outcomes presented in Fig. 1. A statistically more powerful modification of a step-wise between-tiers comparison MBD randomization-test procedure of Levin et al. (2018) that was originally proposed by Revusky (1967) is sensitive to assessing that pattern and yielded statistically significant results ($ps < .023$) on all measures except the fractions test ($p = .17$). Thus, it can be concluded that students' observed

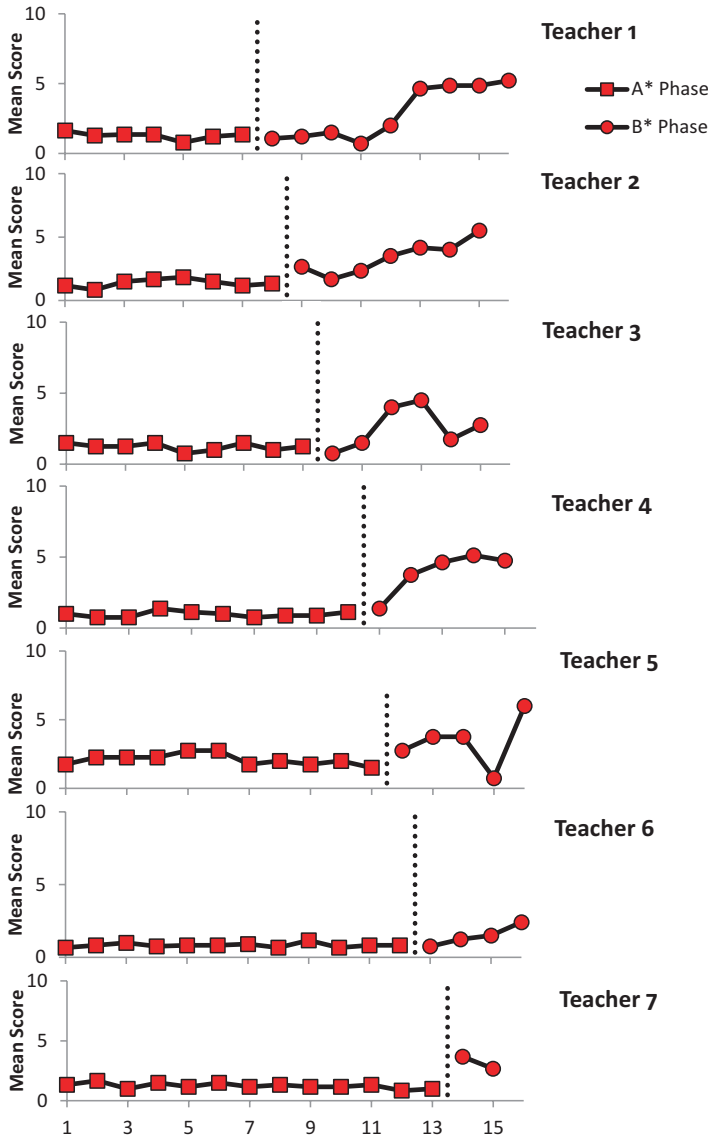


Fig. 1 Quality of mathematical reasoning

mathematics performance increases occurred, at least to some extent, in accord with the hoped-for staggered tier improvements of successful MBD intervention outcomes.

Second, based on the aforementioned visual analyses of the data, it was noted that on two of the outcome measures (Mathematical Reasoning and Argumentative Elements) students exhibited a low stable level of performance during their baseline phase, followed by a steady increase in performance during their intervention phase.

In formal between-phase regression line “slope-change” analyses (Levin et al., 2021), no two-observation delay was included because Class 7 did not provide enough intervention-phase outcomes for a meaningful slope to be calculated. An illustration of the present study’s baseline-to-intervention-phase slope increases may be seen for the Mathematical Reasoning measure presented in Fig. 1, where the actual intervention start point for each class occurred two observations earlier than is indicated there. From a statistical perspective, Wampold-Worsham MB randomization tests documented a significant increase in the slopes of students’ performance between the baseline and intervention phases, from an across-students average slope of $-.01$ to $.52$, $p = .026$, and from $-.02$ to $.70$, $p = .012$, for Mathematical Reasoning and Argumentative Elements, respectively.

3.1 Fine-Grained Analyses: Or with a Tip of the Hat to an Esteemed Mentor, Colleague, and Friend: “Fine Grahamed” Analyses

The preceding analyses were conducted on the comprehensive 20 parallel forms assessments that included all the FACT intervention content. During the earlier parts of most students’ intervention phase, those assessments contained content that had not yet been covered by the classroom teachers. Thus, students’ performance on such content for those assessments would not be expected to be responsive to any intervention effects. A seemingly more appropriate analysis was therefore conducted on just the lesson content that would have been covered for all students on all assessments. As with the mixed results already reported for the complete test content: (a) the Wampold-Worsham MBD test indicated that there was no statistically significant A- to B-phase improvement in students’ performance ($p = .47$); but (b) the Levin et al.’s (2018) Modified Revusky MBD test again produced a statistically significant appropriately differentiated vertical pattern of improvement ($p = .037$).

Pretest and Posttest Measures of Fraction Operations

A correlated-samples t -test was conducted on the distal 27-item fraction operations measure that was administered to the 27 participating students at the beginning and the end of the study. Students exhibited a statistically significant increase on that measure ($p < .001$), amounting to a 3-item gain, $d = 1.30$. For this twice-administered measure, the observed improvement over the course of the school year does not reflect the true impact of the FACT + R²C² intervention *per se* because it is entangled with regular school mathematics instructional content by teachers during the baseline phases and student growth.

4 Discussion

In our single-case classroom intervention investigation, we aimed to understand whether writing-to-learn mathematics through argument writing would increase students' fraction knowledge, their quality of mathematical reasoning, and the number of argumentative elements and total words written when provided with explicit strategies instruction using the six stages of SRSD. Based on previous research showing the benefits of writing-to-learn content (Bangert-Drowns et al., 2004; Graham et al., 2020; Hubner et al., 2006), as well as benefits of explicit writing strategies instruction (Gillespie & Graham, 2014; Graham et al., 2012), we expected that learning to construct written arguments and to critique the reasoning of peers would: (a) help students with MLD address their misconceptions when solving fraction problems and (b) allow those students to develop a deeper understanding of fractions (Hacker et al., 2019; Kiuahara et al., 2020). Our analyses uncovered selected classroom-level effects on all four outcome measures, similar to our findings in the MBD study presented by Hacker et al. (2019). That is, some teachers and certain outcome measures were associated with more student gains than others. This could be attributed to several factors.

First, the FACT intervention using SRSD combines multiple learning components (i.e., fraction learning and argumentative writing) with behavioral components (i.e., self-regulating learning behaviors and self-monitoring affect and motivation), which students acquire. These multiple components may require extended class sessions for teachers to ensure a student with MLD can independently use the strategies. Kiuahara et al. (2020) found that on average and with high treatment fidelity across teachers, the general education teachers in the FACT condition completed the lessons in 23 45-min sessions compared to the special education teachers. The latter completed the lessons in 30 45-min sessions. This indicated that students with MLD may require extended time reaching the Independent Stage of SRSD, as we found in the current study. For the seven teachers in this study, the instructional pacing to complete the lessons ranged from 16 days to 42 days. The teachers in the intervention phase for the longest amount of time demonstrated high treatment fidelity and did not require additional coaching and support; however, they had a large number of ELs with MLD in their classrooms (range 33–57%).

Based on their research using SRSD to teach ELs the process of argumentative writing, Cuenca-Carlino et al. (2018) suggest incorporating culture and specific language needs into SRSD. The quality of teacher-student and student-student interaction during learning influences language development (Gersten et al., 2009; Klingner & Soltero-Gonzales, 2009). For example, although SRSD addresses components of student agency (i.e., students are taught strategies to engage in discourse affecting their learning behaviors), it may also be necessary for teachers to understand how (a) to expand students' ability to draw on their own languages and cultural experiences for reflecting on their original or novel approaches when solving problems (Klinger and Soltero-Gonzales) and (b) to facilitate ways for ELs to think aloud using their first language or with their peers when solving math problems (Garcia &

Sylvan, 2011). Conversely, the teacher who completed the FACT lessons in the shortest amount of time (i.e., 16 days) reported in her teaching log that she did not engage students in any supplemental learning activities and simply followed the manualized protocol for implementing the lessons almost verbatim. Although her treatment fidelity was high (91%), we learned that PD should include space for teachers to discuss the recursive approach to teaching SRSD and decision-making processes for determining when teachers should reteach, modify, or reorder the stages of SRSD (Cuenca-Carlino et al., 2018). Students benefit only when they are provided multiple opportunities to connect mathematical language and communicate their learning (Bangert-Drowns et al., 2004; Klein, 1999).

Second, one of our concerns from the previous study reported by Hacker et al. (2019) was to increase teacher “buy in” to implement a novel instructional intervention such as FACT. One way we addressed this was here to devote time during PD to discuss with teachers their experiences they had teaching fractions to students with MLD and some of the common misconceptions that students have when learning fractions (Borko et al., 2015). We also included a protocol for using questioning as a way for teachers to assess students’ understanding and to provide opportunities for students to articulate their understandings (Borko et al., 2015; Polly et al., 2014; Jayanthi et al., 2017). We found that treatment fidelity was high across all seven teachers. It also proved beneficial for us to have established a minimum criterion level for knowing when to provide teachers with additional coaching during the implementation phase. For example, one teacher who did not meet the criterion of treatment fidelity (*viz.*, greater than 85%) on more than one observed lesson had transferred the learning activities from the teacher’s manual to a SmartBoard that the teacher used daily. This resulted in some of the FACT content not having been easily transferred and therefore was skipped by the teacher during instruction. The third author met with the teacher regularly to address the fidelity components that were missed during the observation and answered any questions the teacher had about the lesson.

Third, from a methodological perspective, our MBD with random assignment of teachers to tiers and associated randomization statistical tests represents a rigorous single-case intervention design (Kratochwill & Levin, 2010; Levin et al., 2019). Although logistical constraints and exigencies rendered it not possible to implement the intervention optimally here, future scientifically credible and statistically powerful single-case research that focuses on teaching SRSD should strive to include a greater stagger of the tiers’ intervention start points, along with statistical randomization-test models that require a start point for each tier member that is randomly sampled from two or more acceptable potential intervention start points (see, for example, Levin & Ferron, 2021).

Finally, we did not account for the type of instruction the students received during the baseline phase, which ranged from 19 to 42 days, depending on the teacher’s staggered intervention start date and the possible overlap with the district’s pacing calendar for teaching or reviewing fractions. Taking all these various design and operational challenges into consideration (Kratochwill et al., 2021) leads directly to a general guiding principle that has emerged from our present and earlier FACT/

SRSD studies (see Hacker et al., 2019; Kiuahara et al., 2020). As with large-scale randomized controlled trials studies, in single-case classroom-based intervention investigations one's mantra should be: "Plan for the best but always expect the unexpected."

In sum, our previous FACT intervention studies found that engaging students with MLD in activities that focused on communicating with precise mathematical language and constructing arguments increased students' quality of mathematical reasoning, argumentative elements, and total words written (Hacker et al., 2019; Kiuahara et al., 2020). Although certainly not conclusive, the collective visual and statistical outcomes from this study are encouraging, in that they are suggestive of positive effects associated with using writing-to-learn mathematics and SRSD. They should provide continued motivation for classroom-based interventionists to extend this line of research in teaching mathematics and writing, especially for a wide range of ELs with MLD.

Acknowledgements We wish to thank the teachers and students who participated in this study, Drs. Noelle Converse and Leah Vorhees for their support, and the Utah State Board of Education Special Education Services for providing the funding for this project.

References

- Bailey, D. H., Zhou, X., Zhang, Y., Cui, J., Fuchs, L., Jordan, N. C., Gersten, R., & Siegler, R. S. (2015). Development of fraction concepts and procedures in U.S. and Chinese children. *Journal of Experimental Psychology*, *129*, 68–83.
- Ball, D. L., & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 27–44). National Council of Teachers of Mathematics.
- Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, *74*, 29–58.
- Borko, H., Jacobs, J., Koellner, K., & Swackhamer, L. E. (2015). *Mathematics professional development: Improving teaching using the problem-solving cycle and leadership preparation models*. National Council of Teachers of Mathematics.
- Boscolo, P., & Mason, L. (2001). Writing to learn, writing to transfer. In G. Rijlaarsdam (Series Ed.), *Studies in writing: Volume 7. Writing as a learning tool: Integrating theory and practice* (pp. 83–104). Kluwer Academic Publishers.
- Bryant, B. R., & Bryant, D. P. (2008). Mathematics and learning disabilities. *Learning Disability Quarterly*, *31*, 3–11.
- Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single case research design and analysis: New directions for psychology and education* (pp. 187–212). Lawrence Erlbaum Associates, Inc.
- Common Core State Standards for Mathematics*. (2010). Retrieved January 1, 2014, from www.corestandards.org
- Cuenca-Carlino, Y., Gozur, M., Jozwik, S., & Krissinger, E. (2018). The impact of self-regulated strategy development on the writing performance of English learners. *Reading and Writing Quarterly*, *34*, 248–262. <https://doi.org/10.1080/10573569.2017.1407977>

- Desoete, A. (2015). Language and math. In P. Aunio, R. Mononen, & A. Laine, *Mathematical learning difficulties—Snapshots of current European research. LUMAT*, 5, 647–674.
- Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C., Cirino, P. T., Jordan, N. C., Siegler, R., Gersten, R., & Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology*, 105, 683–700.
- Gafurov, B. S., & Levin, J. R. (2022). *ExPRT (Excel Package of Randomization Tests): Statistical analyses of single-case intervention data*. Current Version 4.3 (November 2022) is retrievable from the ExPRT website at <http://ex-prt.weebly.com>
- Garcia, O., & Sylvan, C. E. (2011). Pedagogies and practices in multilingual classrooms: Singularities in pluralities. *Modern Language Journal*, 95, 385–400.
- Geary, D. C. (2011). Consequences, characteristics, and causes of mathematical learning disabilities and persistent low achievement in mathematics. *Journal of Developmental & Behavioral Pediatrics*, 32, 250–263.
- Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79(3), 1202–1242. <https://doi.org/10.3102/003454309334431>
- Gillespie, A., & Graham, S. (2014). A meta-analysis of writing interventions for students with learning disabilities. *Exceptional Children*, 80, 454–473.
- Graham, S., McKeown, D., Kiuahara, S. A., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, 104(4), 879–896.
- Graham, S., Kiuahara, S. A., & MacKay, M. (2020). The effects of writing on learning in science, social studies, and mathematics: A meta-analysis. *Review of Educational Research*, 90(2), 179–226. <https://doi.org/10.3102/0034654320914744>
- Hacker, D. J., & Dunlosky, J. (2003). Not all metacognition is created equal. *Problem-based learning for the information age. New Directions for Teaching and Learning*, 95, 73–80.
- Hacker, D. J., Kiuahara, S. A., & Levin, J. R. (2019). A metacognitive intervention for teaching fractions to students with or at-risk for learning disabilities in mathematics. Special issue on metacognition in mathematics education. *ZDM: The International Journal on Mathematics Education*, 51(4), 601–612. <https://doi.org/10.1007/s11858-019-01040-0>
- Harris, K. R., & Graham, S. (2009). Self-regulated strategy development in writing: Premises, evolution, and the future. *Teaching and Learning Writing*, 6, 113–135.
- Harris, K. R., Lane, K. L., Graham, S., Driscoll, S. A., Sandmel, K., Brindle, M., & Schatschneider, C. (2012). Practice-based professional development for self-regulated strategies development in writing: A randomized controlled study. *Journal of Teacher Education*, 63(2), 103–119. <https://doi.org/10.1177/002248711429005>
- Harris, K. R., Graham, S., & Adkins, M. (2014). Practice-based professional development and self-regulated strategy development for Tier 2, at-risk writers in second grade. *Contemporary Educational Psychology*, 40(1), 5–16.
- Hebert, M. A., & Powell, S. R. (2016). Examining fourth-grade mathematics writing: Features of organization, mathematics vocabulary, and mathematical representations. *Reading and Writing*, 29, 1511–1537. <https://doi.org/10.1007/s11145-016-9649-5>
- Horner, R. H., & Odom, S. L. (2014). Constructing single-case research designs: Logic and options. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 27–51). American Psychological Association.
- Hubner, S., Nuckles, M., & Renkl, A. (2006). Prompting cognitive and metacognitive processing in writing-to-learn enhances learning outcomes. In *Proceedings of the 28th annual conference of the cognitive science society* (pp. 357–362). Erlbaum.
- Hughes, E. M., Witzel, B. S., Riccomini, P. J., Fries, K. M., & Kanyongo, G. Y. (2014). A meta-analysis of algebra interventions for learners with disabilities and struggling learners. *Journal of the International Association of Special Education*, 15(1), 36–47.

- Jayanthi, M., Gersten, R., Taylor, M. J., Smolkowski, K., & Dimino, J. (2017). *Impact of the developing mathematical ideas professional development program on grade 4 students' and teachers' understanding of fractions* (REL 2017-256). Regional Educational Laboratory Southeast.
- Jitendra, A. K., & Star, J. R. (2011). Meeting the needs of students with learning disabilities in inclusive mathematics classrooms: The role of schema-based instruction on mathematical problem solving. *Theory Into Practice, 50*, 12–19.
- 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.
- Kiuahara, S. A., O'Neill, R., Hawken, L. S., & Graham, S. (2012). The effectiveness of teaching 10th grade students with a disability STOP, AIMS, and DARE for planning/drafting persuasive text. *Exceptional Children, 78*(3), 335–355.
- Kiuahara, S. A., Kratochwill, T. R., & Pullen, P. C. (2017). Designing robust experimental single-case design research. In J. M. Kauffman, D. P. Hallahan, & P. C. Pullen (Eds.), *Handbook of special education* (2nd ed., pp. 116–136). Routledge.
- Kiuahara, S. A., Gillespie Rouse, A., Dai, T., Witzel, B., Morphy, P., & Unker, B. (2020). Constructing written arguments to develop fraction knowledge. *Journal of Educational Psychology, 112*(3), 584–607. <https://doi.org/10.1037/edu0000391>
- Klein, P. D. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review, 11*, 203–270.
- Klingner, J. K., & Soltero-Gonzalez, L. (2009). Culturally and linguistically responsive literacy instruction for English language learners with learning disabilities. *Multiple Voices for Ethnically Diverse Exceptional Learners, 12*, 4–10.
- Korhonen, J., Linnanmaki, K., & Aunio, P. (2012). Language and mathematical performance: A comparison of lower secondary school students with different level of mathematical skills. *Scandinavian Journal of Educational Research, 56*(3), 333–344. <https://doi.org/10.1080/000313831.2011.599423>
- Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods, 15*, 124–144. <https://doi.org/10.1037/a0017736>
- Kratochwill, T. R., Horner, R. H., Levin, J. R., Machalicek, W., Ferron, J., & Johnson, A. (2021). Single-case design standards: An update and proposed upgrades. *Journal of School Psychology, 89*, 91–205.
- Krowka, S. K., & Fuchs, L. S. (2017). Cognitive profiles associated with responsiveness to fraction intervention. *Learning Disabilities Research & Practice, 32*, 216–230.
- Levin, J. R. (1992). Single-case research design and analysis: Comments and concerns. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 213–224). Erlbaum.
- Levin, J. R., & Ferron, J. M. (2021). Different randomized multiple-baseline models for different situations: A practical guide for single-case intervention researchers. *Journal of School Psychology, 86*, 169–177.
- Levin, J. R., Ferron, J. M., & Gafurov, B. S. (2017). Additional comparisons of randomization-test procedures for single-case multiple-baseline designs: Alternative effect types. *Journal of School Psychology, 63*, 13–34.
- Levin, J. R., Ferron, J. M., & Gafurov, B. S. (2018). Comparison of randomization-test procedures for single-case multiple-baseline designs. *Developmental Neurorehabilitation, 21*, 290–311.
- Levin, J. R., Kratochwill, T. R., & Ferron, J. M. (2019). Randomization procedures in single-case intervention research contexts: (Some of) “the rest of the story”. *Journal of the Experimental Analysis of Behavior, 112*, 334–348.
- Levin, J. R., Ferron, J. M., & Gafurov, B. S. (2021). Investigation of single-case multiple-baseline randomization tests of trend and variability. *Educational Psychology Review, 33*, 713–737.

- Mazzocco, M. M. M., Myers, G. F., Lewis, K. E., Hanich, L. B., & Murphy, M. M. (2013). Limited knowledge of fraction representations differentiates middle school student with mathematics learning disability (dyscalculia) versus low math achievement. *Journal of Experimental Child Psychology, 115*, 371–387.
- National Center for Education Statistics, Institute of Education Sciences, National Assessment of Educational Progress for Mathematics. (2015). *The nation's report card: Mathematics 2015*. Retrieved from <https://nces.ed.gov/nationsreportcard/mathematics/>
- National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. U.S. Department of Education. Retrieved from www2.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf
- Newell, G. E. (2006). Writing to learn: How alternative theories of school writing account for student performance. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 235–247). Guilford Press.
- Parker, R. I., Vannest, K. J., & Davis, J. L. (2014). Non-overlap analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 127–151). American Psychological Association.
- Polly, D., Neale, H., & Puglee, D. K. (2014). How does ongoing task-focused mathematics professional development influence elementary school teachers' knowledge, beliefs and enacted pedagogies? *Early Childhood Educational Journal, 42*, 1–10.
- Resnick, L. B. (1987). *Education and learning to think*. National Academy Press.
- Revusky, S. H. (1967). Some statistical treatments compatible with individual organism methodology. *Journal of the Experimental Analysis of Behavior, 10*, 319–330.
- 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.
- Slocum, T. A., Pinkelman, S. E., Joslyn, P. R., & Nicols, B. (2022). Threats to internal validity in multiple-baseline design variations. *Perspectives on Behavior Science*. <https://doi.org/10.1007/s40614-00326-1>
- Tindal, G., & Alonzo, J. (2012). *easyCBM*. Houghton Mifflin Harcourt.
- Vukovic, R. K., & Lesaux, N. K. (2013). The language of mathematics: Investigating the ways language counts for children's mathematical development. *Journal of Experimental Child Psychology, 115*, 227–244.
- Wampold, B., & Worsham, N. (1986). Randomization tests for multiple-baseline designs. *Behavioral Assessment, 8*, 135–143.
- Wechsler, D. (2009). *Wechsler individual achievement test* (3rd ed.). Pearson.
- Wilkinson, G. S., & Robertson, G. J. (2006). *Wide range achievement test* (4th ed.). PAR.
- Witzel, B. S., Mercer, C. D., & Miller, M. D. (2003). Teaching algebra to students with learning difficulties: An investigation of an explicit instruction model. *Learning Disabilities Research & Practice, 18*(2), 121–131.
- Woodward, J., Beckmann, S., Driscoll, M., Franke, M., Herzig, P., Jitendra, A., Koedinger, K. R., & Ogbuehi, P. (2012). *Improving mathematical problem solving in grades 4 through 8: A practice guide* (NCEE 2012-4055). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications_reviews.aspx#pubsearch/