Comparing and Discussing Multiple Strategies: An Approach to Improving Algebra Instruction

Kelley Durkin[1], Bethany Rittle-Johnson[2], Jon R. Star[3], & Abbey Loehr[4]

[1]Department of Teaching and Learning, Vanderbilt University

[2]Department of Psychology and Human Development, Vanderbilt University

[3]Graduate School of Education, Harvard University

[4]Institute for School Partnership, Washington University in St. Louis

Author Note

[1]Correspondence concerning this article should be addressed to Kelley Durkin, 230 Appleton Place, PMB 181, Vanderbilt University, Nashville, TN 37203, United States of America. Email: kelley.durkin@vanderbilt.edu

This work was supported by the National Science Foundation under grant DRL 1561286. The ideas in this paper are those of the authors and do not represent official positions of NSF. We would like to thank Stephanie Fine, Jessica Sommer, Mason Shero, Shelby Jacobsen, Ranique Jackson, Molly Sullivan, Kathy Boehm, Xueru Yu, Yi Zhou, Gabrielle Rademaker, Tykeena Watson, and Allison Yi for their help with preparing and implementing the study. We would also like to thank the teachers and students who participated in this project.

Citation

Durkin, K., Rittle-Johnson, B., Star, J. R., & Loehr, A. (in press). Comparing and discussing multiple strategies: An approach to improving algebra instruction. *The Journal of Experimental Education*. doi: 10.1080/00220973.2021.1903377

Abstract

Productive learning of algebra is supported when students reflect on multiple strategies, compare them and discuss the rationale behind and relative merits of particular strategies. ***Comparison and Discussion of Multiple Strategies (CDMS)*** is an instructional approach designed to support these processes in math classrooms**.** In the current study, 16 Algebra I teachers received professional development and supplemental materials to support CDMS when teaching a unit on linear equation solving and 475 of their students completed assessments of their linear equation solving knowledge before and after the unit. Thirteen Algebra I teachers and their 359 students were the business-as-usual control group. CDMS increased how often teachers engaged their students in comparison of multiple strategies, sustained small group work, and sustained mathematical discussions. Students in CDMS classrooms also had higher knowledge of linear equations on the posttest, particularly procedural flexibility, even after controlling for pretest knowledge and school demographic differences. Thus, encouraging teachers to regularly compare and discuss multiple strategies increases students' algebra learning. Findings highlight the need to expand theories of algebra learning to include attention to procedural flexibility, illustrate an instructional theory and method to promote broader learning about algebra, and provide evidence for effective instructional practices.

Keywords: mathematics education; procedural flexibility; comparison; multiple strategies; discussion; classroom intervention

Comparing and Discussing Multiple Strategies: An Approach to Improving Algebra Instruction

Research and policy suggest that proficiency in algebra is critical to academic, economic, and life success. For example, success in algebra is necessary for access to higher mathematics and is correlated with positive life outcomes, such as college graduation (Adelman, 2006; National Mathematics Advisory Panel, 2008). Unfortunately, national and international assessments have drawn attention to pervasive student difficulties in algebra (Beaton et al., 1996; Blume & Heckman, 1997; Lindquist, 1989; National Center for Education Statistics, 2020a; Schmidt et al., 1999). For example, only 32% of 15-year-old students worldwide were at Level 4 or above (out of 6 levels) on the algebra subscale of the Programme for International Student Assessment (PISA), having some flexibility in interpreting and reasoning about functional relationships (OECD, 2013). Difficulties with algebra are particularly prevalent among students who are not educated in East Asia. For example, students in the United States continue to struggle on very straightforward algebra problems; only 36% of eighth graders could interpret the meaning of a linear equation in context correctly, and only 22% of eighth graders could convert degrees Fahrenheit into degrees Celsius using a linear equation (National Center for Education Statistics, 2020b).

There are several reasons algebra can be difficult for students, one of which is because it is substantially different from arithmetic (Kilpatrick et al., 2001). In contrast to arithmetic's focus on concrete and countable objects, algebra is the first time where students engage in prolonged abstraction and symbolization (Kieran, 1992). Algebra is fundamentally concerned with generalizing and expressing relationships among quantities, often using symbols (Kieran, 1992; Lloyd et al., 2011). In addition, expressing and analyzing these relationships between quantities requires facility not only with symbols, but also with multiple representations, including tables and graphs (Lloyd et al., 2011). Success in algebra requires an understanding of key mathematical concepts as well as flexibility using a variety of simple and complex symbolic strategies (Star & Newton, 2009). We theorize that productive learning of algebra is supported by reflection on multiple strategies through comparing them and discussing the rationale behind and relative merits of particular strategies.

Developing Algebraic Knowledge

Proficiency in algebra requires developing conceptual knowledge, procedural knowledge, and procedural flexibility. Conceptual knowledge is the knowledge of abstract, general principles (Byrnes & Wasik, 1991; Canobi, 2009; Rittle-Johnson et al., 2001). Conceptual knowledge is of critical importance; understanding of key concepts such as equivalence and variable is essential to success in algebra (Knuth et al., 2006). Procedural knowledge is the knowledge of mathematical strategies and often develops through problem-solving practice, and thus is tied to particular problem types (Baroody et al., 2007; Byrnes & Wasik, 1991; Canobi, 2009; Rittle-Johnson et al., 2001).

Students also need to know multiple strategies for solving problems and be able to select the most appropriate strategy for a given problem, often called procedural flexibility (Star, 2005; Star & Seifert, 2006; Verschaffel et al., 2009). Learners who develop procedural flexibility are more likely to use or adapt existing strategies when faced with unfamiliar problems and to have greater conceptual knowledge (Blöte et al., 2001; Hiebert et al., 1996). Furthermore, procedural flexibility is a salient characteristic of experts in mathematics (Dowker, 1992; Star & Newton, 2009). Procedural flexibility is distinct from, but related to, conceptual and procedural knowledge (Schneider et al., 2011). For example, in a sample of middle school students with some prior knowledge of algebra, the latent variables for each knowledge type were correlated .63-.66, and model comparisons confirmed that the data were better represented using three distinct knowledge types rather than a single type. Further, conceptual and procedural knowledge at the beginning of the unit contributed independently to developing procedural flexibility.

Potential Benefits of Multiple Strategies, Comparison, and Discussion in Algebra

Comparison and Discussion of Multiple Strategies (***CDMS***) is an approach to more productively support learning of algebra, including procedural flexibility, in algebra instruction. It is based on converging evidence from cognitive science and mathematics education research on the importance of multiple strategies, comparison, and discussion for mathematics learning.

First, knowledge of multiple strategies is beneficial for students' mathematics learning. Knowing diverse strategies allows people to adapt their strategy use to task demands. For example, when solving arithmetic problems, children usually select faster and less effortful strategies on easier problems and slower and more effortful strategies on more difficult ones (Siegler, 1996). Knowing multiple strategies helps learners to adapt to new problem features and supports better understanding of the domain (Blöte et al., 2001; Hiebert et al., 1996). Silver and colleagues (2005) note, "It is nearly axiomatic among those interested in mathematical problem solving as a key aspect of school mathematics that students should have experiences in which they solve problems in more than one way" (p. 288).

In elementary school mathematics, high-quality instruction often focuses on strategies invented by students (Carpenter et al., 1989). However, this practice is less common in algebra classrooms. Students are rarely encouraged to invent strategies for solving symbolic algebra problems. Even when they are encouraged to invent strategies, the invented strategies are rarely productive, and the practice does not promote procedural flexibility (Authors, 2008).

Furthermore, it is not clear whether most students could invent strategies for some algebra topics such as quadratics. Rather, demonstrating alternative, more efficient strategies in writing through worked examples is an effective way to support procedural flexibility (Authors, 2008). Worked examples present the solution steps for solving a problem, and decades of research indicates that including worked examples in instruction improves learning (e.g., Atkinson et al., 2000).

Second, comparison of multiple strategies is a powerful process that supports learning, including algebra learning. Comparison improves learning in many domains, ranging from preschoolers learning new words (Namy & Gentner, 2002) to business school students learning contract negotiation strategies (Gentner et al., 2003). A meta-analysis confirmed that comparison promotes learning across a range of domains (Alfieri et al., 2013). There is rigorous evidence from short-term classroom studies on the benefits of comparison for mathematics learning in particular. Across five short-term classroom studies, having learners compare multiple strategies led to greater mathematics knowledge than studying the same strategies sequentially, without comparison (Authors, 2007, 2009a, 2009b, 2009c, 2012). For example, in a one-week experimental study, Authors (2007) randomly assigned 7th grade student pairs to learn multistep linear equation solving by either comparing worked examples of two different strategies for solving the same problem presented side-by-side, or by studying the worked examples of the strategies one at a time (sequentially). Students in the comparison condition demonstrated greater procedural knowledge and procedural flexibility than students in the sequential condition. Comparison has also been shown to improve conceptual knowledge of algebra (Authors 2009b, 2009c), estimation knowledge (Authors, 2009a) and knowledge about the altitude of a triangle (Guo & Pang, 2011). Further, Practice Guides from the U.S. Department of Education identified comparing multiple strategies as an evidence-based recommendation for improving mathematical problem solving in the middle grades and choosing from alternative algebraic strategies when solving problems as a recommendation for improving algebra knowledge (Star et al., 2015; Woodward et al., 2012). However, comparing multiple strategies can overwhelm students with low prior knowledge and reduce learning relative to studying multiple strategies one at a time (Authors, 2009). In particular, it is easier for students to learn from comparing an unfamiliar strategy to a familiar strategy, such as comparing a new strategy to a strategy students have already learned. Students can learn from comparing two unfamiliar strategies, but it requires additional support, such as providing more time to understand and compare the strategies and providing carefully-crafted explanation prompts that guide students' attention towards key ideas (Authors, 2012).

Comparing multiple strategies is also integral to "best practices" in mathematics education. Having students share and compare strategies for solving a particular problem lies at the core of reform pedagogy in many countries throughout the world (Australian Education Ministers, 2006; Kultusministerkonferenz, 2004; National Council of Teachers of Mathematics, 2014; Singapore Ministry of Education, 2006; Treffers, 1991). This includes the Common Core Standards in mathematics in the U.S. (*Common Core State Standards in Mathematics*, 2010). Indeed, expert teachers in the U.S and East Asia have students compare multiple strategies (Ball, 1993; Lampert, 1990; Richland et al., 2007).

Third, the development of mathematics knowledge is believed to be enhanced by classroom discussions in which students generate explanations and teachers facilitate a discussion of different student responses (Lampert, 1990; Silver et al., 2005; Stein et al., 2008). Indeed, prompting people to discuss new information improves learning across a variety of mathematics topics and age groups (Aleven & Koedinger, 2002; Hodds et al., 2014; Rittle-

Johnson, 2006) and more frequent engagement with other students' strategies and ideas during discussion is related to greater success on a mathematics assessment (Webb et al., 2014). Teachers use several techniques to support these discussions. For example, 3rd and 4th grade teachers asked students to compare their own ideas to other students' ideas and used specific language to encourage students to monitor their own and each other's ideas and understanding (Webb et al., 2014). Overall, discussion and explanation support knowledge integration and knowledge generalization (Chi, 2000; Webb et al., 2014).

Development of Materials to Support Comparison and Discussion of Multiple Strategies

To promote high-quality adoption in classrooms, creating curriculum materials and teacher professional development (PD) are frequently needed. Teachers in the U.S. often struggle to effectively support and discuss comparison of multiple strategies (Stein et al., 2008). In fact, analyses of video records of mathematics instruction indicate that comparison is often not well enacted in U.S. classrooms (Richland et al., 2004, 2007), including in algebra instruction (Litke, 2020). It can be difficult for teachers to spontaneously generate the most effective explanation prompts for students. For instance, Asian teachers often ask their students to explain the most challenging aspects of comparison, whereas U.S. teachers often ask their students to explain the simple aspects (e.g., identifying a similarity) but explain the more difficult components themselves (Hiebert et al., 2003; Richland et al., 2012).

To engage students in comparison and discussion of multiple strategies, Stein and colleagues (2008) suggest that teachers need to anticipate, monitor, select, sequence, and make connections between students' strategies and responses. Due to the difficulties many teachers face in effectively implementing these key instructional practices, CDMS instructional materials were developed that include worked examples that anticipate common student strategies, selected to capture important differences in strategies and sequenced and paired to facilitate meaningful comparison. A team of mathematics education experts, including researchers, mathematicians, and Algebra I teachers, developed the materials by going through a typical Algebra I course syllabus, identifying core concepts, common student difficulties, and key misconceptions, and then creating materials to attempt to address them. The intent of the materials is to scaffold teachers' effective use of comparison.

At the core of this supplemental Algebra I curriculum were worked example pairs (WEPs). Each WEP showed the mathematical work and dialogue of two hypothetical students as they attempted to solve one or more algebra problems. The curriculum contained four types of WEPs, with the types varying in what is being compared and the instructional goal of the comparison. First, *Which is better?* WEPs showed the same problem solved using two different, correct strategies, with the goal of understanding when and why one strategy is more efficient or easier than another strategy for a given problem (e.g., Authors, 2007). Second, *Which is correct?* WEPs showed the same problem solved with a correct and incorrect strategy, with the goal of understanding and avoiding common errors (Durkin & Rittle-Johnson, 2012). The curriculum also included two additional types of WEPs, which first emerged during a classroom study by Newton et al. (2010) and were further developed by mathematics educators on the research team. *Why does it work?* WEPs showed the same problem solved with two different correct strategies, but with the goal of illuminating the conceptual rationale in one strategy that is less apparent in the other strategy. *How do they differ?* WEPs showed two different problems solved in related ways, with an interest in illustrating what the relationship between problems and answers of the two problems revealed about an underlying mathematical concept. A series of increasingly sophisticated discussion prompts promoted specific comparisons and generalizations tailored to the learning goal of the lesson. Professional development also provided training and practice promoting student explanation and discussion.

This work resulted in a yearlong randomized controlled trial that explored the feasibility of implementation of the Algebra I supplemental curriculum and its impact on teachers' instruction and students' mathematical knowledge (Authors, 2015b). Observations, surveys and interviews indicated that the professional development was successful in familiarizing teachers with the CDMS approach (Authors, 2014a), that teachers were able to implement the CDMS materials with reasonable fidelity (Authors, 2015b), and that students enjoyed and found valuable the emphasis on multiple strategies (Authors, 2014b). However, the results of the randomized controlled trial indicated that there was no main effect of condition on student achievement, in large part because use of the supplemental curriculum was infrequent by many teachers; greater use of the curriculum did predict greater student learning (Authors, 2015b).

Current Study

Building on this past work, we revised the CDMS materials to increase guidance for teachers on how to integrate the WEPs and discussion into their lessons, more explicitly emphasize the role discussion plays in supporting effective comparison, improve lesson closure, and focus more on particular algebra units. To increase the frequency of use of the materials, we aligned materials more closely with teachers' curriculum materials (e.g., linked WEPs to specific lessons and chapters), provided more specific guidance about when in the lesson a WEP could be used, and during professional development, we helped teachers plan when to use what materials. The current study focused on the effectiveness of our first unit, which was on linear equation solving. We investigated how our revised materials and professional development on CDMS affected teachers' instructional practice and students' learning of linear equation solving compared to typical algebra instruction.

Method

Participants

As part of a delayed treatment design, in 2018-2019, 16 Algebra I teachers across four schools in four districts in Massachusetts and New Hampshire participated in the treatment group, including 7 treatment teachers who had piloted the intervention the previous school year. For the business-as-usual control group, 13 Algebra I teachers across six schools in four districts in Massachusetts and New Hampshire participated (and received the treatment PD and materials the following year). Assignment to condition occurred at the school level, and schools were assigned to the treatment condition until we met our target number of teachers. Two teachers taught 8th grade (1 treatment and 1 control) and the remaining teachers taught 9th grade. Based on a teacher survey, the treatment and control teachers were experienced teachers (average years of teaching experience = 11.1 (range 1-28 years) and 13.5 (range 3-40 years), respectively), were experienced teachers of algebra (average years teaching an algebra course = 9.4 (range 1-28 years) and 8.9 (range 3-20 years), respectively), and had majored in math during undergraduate and/or graduate school (94% and 85% of teachers, respectively).

Classes covered a range of student ability levels. Some teachers taught multiple sections of Algebra I, resulting in the 16 treatment teachers covering 25 sections with a total of 573 students and the 13 control teachers covering 21 sections with a total of 485 students. Seventy-three treatment students and 94 control students did not take the pretest because they were absent the day it was administered. Of the remaining students, 25 treatment students and 32 control students did not take the posttest because they were absent. The analyses for the current study used complete cases with the final analytic sample including 475 students in the treatment group and 359 students in the control group.

Attempts were made to match schools in the treatment and control group on key demographics when recruiting teachers, but the two groups differed in several important ways. The treatment schools had an average of 17% of students receiving free and reduced price lunch (range 6-39), 5% of students were African American or Black (range 1-16), 6% were Hispanic (range 3-14), and 77% were White (range 50-90). The control schools had an average of 35% of students receiving free and reduced price lunch (range 10-47), 6% of students were African American or Black (range 4-14), 26% were Hispanic (range 6-45), and 57% were White (range 32-75). Student-level demographic information was not available. The analytic models described below included these school-level variables as covariates to control for potential differences.

Instructional Materials

We provided treatment teachers with worked example pairs and prompts that encouraged comparison and discussion that aligned with five units in their Algebra I textbooks: Solving Linear Equations, Functions and Graphing Linear Equations, Solving Systems of Linear Equations, Polynomials and Factoring, and Solving Quadratic Equations. The current study focuses on the Solving Linear Equations unit. For this unit, teachers were provided with 9 worked example pairs (*WEP*s), with a sample WEP in Figure 1 (all materials posted on OSF and our website will be shared when unblinded). In this unit, there were 3 *Which is correct?*, 3 *Which is better?*, 2 *Why do they work?* and 1 *How do they differ?* comparisons. The WEPs were similar to those used in past research and showed the work of two hypothetical students who solved a problem followed by prompts for explanation (Authors, 2015a, 2015b). The prompts were designed to help students make comparisons and scaffold discussion, and for each WEP, students were provided with a graphic organizer to support think-pair-share work during the lesson. For each WEP, we also provided a Big Idea page where the hypothetical students identified the learning goal. Each Big Idea page was designed to provide an explicit summary of the WEP's instructional goal because past research suggests that direct instruction helps support student comparisons (Schwartz & Bransford, 1998). For each WEP, teachers were also provided with a teacher guide that included additional prompts and sample student responses.

Treatment teachers participated in professional development (PD) for one week during the summer prior to when they started using our materials that was designed and delivered by the research team, based on our previous professional development (Authors, 2013). The weeklong (35 hours) summer PD introduced teachers to our materials and implementation model (Figure 2). Teachers were given the opportunity to read the materials, view videotaped exemplars of other teachers using the materials, and plan and teach sample lessons using the materials to their peers. Furthermore, teachers evaluated their own and their peers' sample lessons for adherence to the implementation model, using the instrument designed to assess implementation fidelity. They also received feedback on their sample lessons from the research team and their colleagues.

Measures

Assessment. We developed a unit assessment to measure students' conceptual knowledge (e.g., finding a like term), procedural knowledge (e.g., how to solve a linear equation), and procedural flexibility (e.g., selecting the best way to start a problem) based on past work (Authors, 2007, 2009, 2012, 2015). See Table 1 for sample items. Students completed this assessment as a pretest before the unit and as a posttest after the unit. Internal consistency was acceptable at posttest across the 16 items ($\alpha = .75$), but fairly low at pretest ($\alpha = .60$).

Student Handout. We collected completed student handouts (i.e., the student graphic organizer for the think-pair-share routine) to determine which of our materials were used in each class section, as a measure of dosage.

General Fidelity. Teachers in the treatment and control groups were videotaped 2-3 times during each target unit during scheduled visits. For treatment teachers, all recorded lessons intentionally included our materials. Two members of our research team coded 2 randomly selected videos for each teacher for this unit. This coding focused on whether teachers in both groups were using instructional practices emphasized in our implementation model, including comparing multiple strategies for at least 1.5 minutes, engaging in partner/small group work for at least 1 minute, and having a continuous whole class discussion for at least 1.5 minutes. These time durations were chosen because they provided a very liberal estimate of teachers' adherence to our implementation model, as they were half the time requested in the implementation model. In addition, coding of videos from previous implementation years suggested them as sufficient time lengths to capture whether there were extended responses from the teacher and the students. For control lessons, we also coded whether students were exposed to multiple strategies, and if so, whether the multiple strategies were presented side-by-side (this occurred by default when using our materials in the treatment classrooms). The detailed coding scheme is presented in Appendix A in the supplemental materials. Twenty percent of videos were coded by both coders, and reliability across all codes was very high (mean Kappa = .93, range .82 to 1).

Treatment Quality. One randomly selected video from each treatment teacher was also coded for the quality of implementation when using our materials. These codes included Making Sense of Procedures, Supporting Procedural Flexibility, Teaching Questioning, Student Responses, and Opportunities for Interaction, and all codes are described in Table 3 (see

Appendix B in supplemental materials for the detailed coding scheme). The first two codes were adapted from an existing coding scheme for algebra instruction (Litke, 2020), and the code on Teacher Questioning was adapted from Authors (2015a). Each of these five areas was coded with a rating from 1 to 4, with 1 being the lowest (e.g., for the Teacher Questioning code, the teacher was not asking questions or was asking rhetorical questions) and 4 being the highest (e.g., the teacher was asking a significant number of open-ended questions where students were asked to elaborate, to speak for more than one sentence, and to make interpretations). Two experienced mathematics teachers (who had not participated in the project) coded each video and any disagreements were discussed and resolved by the coders. Coders also recorded for how long teachers used our materials within the lesson.

Data Analysis

Data files are posted on OSF. We used multilevel regression models to investigate the effects of instructional practices and condition on students' posttest scores. Students were nested within class sections (ICCs ranging from 0.09 to 0.14), which were nested within schools (ICCs ranging from 0 to 0.04). The main outcome of interest was posttest score on the unit test and the main predictor of interest was assignment to condition. Student unit pretest score, percentage of students at the school receiving free or reduced priced lunch, percentage of African American students at the school, and percentage of Hispanic students at the school were included as covariates. Three similar separate exploratory models were run with posttest subscore on the unit test as the outcome (for conceptual knowledge, procedural knowledge, and flexibility). We also ran models without condition and with an instructional practice central to our implementation model as a predictor to investigate the effects of those practices on students' knowledge.

Results

Instructional Practices

On average, treatment teachers used 7 of the 9 WEPs (range 6-9) and spent about 18 minutes on each WEP (range 13-35 minutes). By design, treatment students were exposed to two strategies presented side-by-side in each of these lessons. In contrast, based on analysis of classroom videos, control teachers rarely exposed their students to multiple strategies and were even less likely to present the multiple strategies side-by-side (see Table 2). When using our materials, treatment teachers usually had sustained comparison of multiple strategies, small group work, and whole class discussions, yet such practices were rare in control classrooms (see Table 2). When treatment teachers were not coded as using an instructional practice, it was usually because they did not implement the instructional practice as long as requested. This is also the reason that so few control teachers' videos were identified as including small group work or whole class discussions – the practice occurred only very briefly. Looking across conditions, students with teachers who more frequently exposed them to multiple strategies and more frequently engaged them in whole class discussion for at least 1.5 minutes performed better on the overall posttest, even after controlling for pretest scores and demographics, $B = 1.81$, $t(39) = 3.08$, $p = .004$ and $B = 1.75$, $t(39) = 2.94$, $p = .005$, respectively. Exposing students to multiple strategies specifically related to higher conceptual knowledge and procedural flexibility at posttest, $B = 0.49$, $t(42) = 2.07$, $p = .045$ and $B = 0.99$, $t(39) = 2.98$, $p = .005$, respectively. Engaging students in whole class discussion also related to higher conceptual knowledge and procedural flexibility at posttest, $B = 0.58$, $t(44) = 2.60$, $p = .013$ and $B = 1.17$, $t(39) = 3.53$, $p = .001$, respectively. None of the other fidelity codes mentioned above significantly predicted students' overall posttest scores (p's > .10).

Coding also captured the quality of treatment teachers' implementation of our approach when using our materials. On a rating scale from 1 to 4, with 4 being the highest quality, teachers on average scored highest on Supporting Procedural Flexibility and Teacher Questioning, and lowest on Opportunities for Interaction (Table 3). With the exception of not providing frequent opportunities for interaction, teachers generally had moderate implementation quality lessons as defined by our coding, with mean scores ranging from 2.56 to 3.09. Analyses of whether treatment teachers' implementation quality predicted students' posttest scores were not appropriate given that only one video per teacher was coded for this unit, with variability in what lesson was being taught and what type of WEP was being used across teachers.

Consider examples of fairly high-quality implementation. Sometimes, treatment teachers asked many conceptual and open-ended questions to help students make sense of procedures. For example, when using a *Why does it work?* WEP where the hypothetical students (Emma and Layla) solved a linear equation using two different strategies, one teacher posed numerous questions, including: “How would you know that's something she is allowed to do?”, “What operation is hiding there when a number is connected to a parenthesis?”, “What does x+1 together mean?”, “Why do both methods get the same answer?”, “Where would you end up with bigger numbers?”, and “What makes Emma's way work?.” Other times, they used questioning to support procedural flexibility. For example, on a *Which is better?* WEP, a teacher asked questions such as "Why would that be effective?", "Which one do you think will be more likely to make mistakes?", "Any other reasons you think Emma's method is better?" and "In this case, which way would you use and why?" In both of these examples, teachers' use of open-ended and conceptual questions resulted in discussions that included both short and extended student responses. In addition, some teachers were successful in getting many students in the classroom to provide responses, not relying solely on a few students to answer questions. And although opportunities for interaction between students were rarely a major focus of the lessons that we analyzed, some teachers created opportunities for interactions by asking a student to rephrase what another student had said, but using their own words (e.g., “Who can rephrase it?”).

Assessments

At pretest, there was not a significant difference between conditions after controlling for covariates, $B = 0.09$, $t(42) = 0.13$, $p = .899$ (see Figure 3). At posttest, students in the treatment condition had higher total scores than students in the control condition (see Table 4 and Figure 3). This difference was mainly due to treatment students having higher procedural flexibility at posttest, although results for the subscales must be interpreted with caution due to the small number of items in each subscale and their low reliability. Nevertheless, they provide important descriptive information of what might be driving the condition differences.

Discussion

Our curriculum and PD supported teachers in engaging their students in comparison of multiple strategies, small group work, and mathematical discussions more frequently than would have happened otherwise. Our intervention also resulted in higher student knowledge of linear equations, particularly procedural flexibility, even after controlling for pretest knowledge and school demographic differences. The findings from this unit are promising; encouraging teachers to compare and discuss multiple strategies can increase students’ knowledge of linear equation solving significantly compared to learning in traditional algebra classrooms. The current study highlights the need to expand theories of algebra learning to include attention to procedural flexibility, illustrates an instructional theory and method to promote broader learning about linear equations, and provides evidence for effective instructional practices, with accompanying coding schemes for systematically capturing important qualities of secondary math instruction.

Expanding Theories of Algebra Learning

Theories of algebra learning usually focus on developing conceptual knowledge (Kieran, 1992). Recent theories also stress connections between multiple representations as a way to build foundational knowledge (e.g., Kaput, 1998), introduction of algebraic concepts into elementary school (Blanton & Kaput, 2005; Carraher et al., 2008), and exploration of real-life contexts, especially when presented in cognitively demanding tasks (Boston & Smith, 2009; Stein et al., 1996). These theories are good but incomplete; they lack an explicit focus on learning symbolic strategies. Complementing knowledge of key concepts, success in algebra requires flexibility in the use of symbolic strategies. The current study illustrates the value of supporting procedural flexibility as well as conceptual and procedural knowledge for algebra learning. Theories of algebra learning need to be expanded to consider the development of procedural flexibility. Comparing and discussing multiple strategies is one way to develop procedural flexibility.

Comparison and Discussion of Multiple Strategies to Promote Algebra Learning

CDMS is a promising instructional theory and approach to promote broad algebra learning. One key component is considering multiple strategies for solving the same problem. Knowing diverse strategies allows people to adapt their strategy use to task demands and to new problem features and supports better understanding of the domain (Blöte et al., 2001; Hiebert et al., 1996; Siegler, 1996). In mathematics education reforms, a common push is for individual students to solve problems in more than one way (Silver et al., 2005). However, generating multiple meaningfully different strategies can be challenging for solving algebraic equations and encouraging students to invent multiple algebraic strategies may not promote procedural flexibility (Authors, 2008). In control classrooms in the current study, multiple strategies for solving the same problem were rarely considered. Presenting two worked examples of how hypothetical students solved a problem is a promising way to have students consider multiple strategies, and treatment teachers readily implemented sustained comparison of the two strategies. We varied how the two strategies differed to promote attention to different dimensions of the strategies, such as accuracy, ease of understanding, and efficiency. Observations indicated that support for procedural flexibility was high in the treatment lessons. Greater knowledge, including procedural flexibility, in students in treatment classrooms compared to those in control classrooms supports this inference, as does the positive relation between frequency of exposure to multiple strategies and posttest scores. Further, findings indicate that discussing and comparing multiple strategies did not confuse students, as confusion would have led to lower procedural knowledge in the treatment condition.

Teachers need some supports in learning how to engage students in strategy comparison. Comparison of multiple strategies for a continuous period was never observed in the control classrooms, and presentation of multiple strategies for solving the same problem was rare (4% of lessons). Past analysis of ninth-grade algebra instruction in the U.S. also indicates that support for procedural flexibility is infrequent, occurring in about one quarter of lessons (Litke, 2020). Future research is needed to explore whether providing materials that illustrate two strategies for solving the same problem may be a useful scaffold for teachers and students to generate and reflect on multiple strategies on their own, without support from specialized materials.

Sustained, whole class discussions are also important to promoting mathematics learning (Lampert, 1990; Silver et al., 2005; Stein et al., 2008). Past research has reported that more frequent engagement with other students' strategies and ideas during discussions in 3rd and 4th grade classrooms was related to greater success on a mathematics assessment and that teachers used a variety of instructional moves to support student engagement (Webb et al., 2014). Many high school mathematics teachers need supports in learning how to lead sustained discussions. Sustained discussions were infrequent in control classrooms in the current study. In a prior attempt to promote a CDMS approach, we also found that treatment teachers had difficulty with the discussion component, with teachers doing most of the explaining (Authors, 2015b). For the current study, we introduced a new think-pair-share instructional routine to promote both small group work and whole class discussion, which seemed to be quite effective at increasing the prevalence and duration of both. As one treatment teacher commented, "And, just the engagement of the discussion, I think, is uncomfortable in the beginning, but once you get used to it, and you start delivering it, it becomes a real powerful tool." Indeed, we found that more frequent whole class discussions were related to greater student success on our algebra assessment, particularly for conceptual knowledge and procedural flexibility. Further, our curriculum was fairly effective at promoting moderate levels of teacher questioning and student responding. Teachers asked open-ended questions, often drawing on the questions we provided and sometimes adding their own follow-up prompts. In response to teacher questioning, students gave a mix of short and extended responses, with some instances of students providing detailed explanations. We had less success supporting treatment teachers in creating a classroom environment where students engaged in mathematical talk with each other and not only with the teacher. Treatment teachers did implement some of the moves for promoting student interaction that they were exposed to in the professional development, including asking multiple students the same question and prompting students to re-voice other students' contributions. However, they rarely integrated questioning with prompts for promoting interaction in ways that encouraged students to listen to, interact with, and respond to each other. Such teaching moves are difficult to directly include in instructional materials. Our PD was not sufficient for learning these techniques, which require responding in real time to students. Our coding schemes for features of instructional quality should aid future research and PD.

The role of sustained small group work in promoting student learning was less clear in the current study. Working with peers in small groups often improves learning, in part because it provides a natural context for encouraging students to generate explanations and provide feedback to one another (Johnson & Johnson, 1994; Webb, 1991). Sustained small group work occurred more often in control classrooms than other features of our intervention, although it was much more common in treatment than control classrooms. Frequency of sustained small group work was not associated with knowledge at posttest. However, limited variability in our measure of small group work and inability to measure quality of the work restricted our ability to detect a relationship. In addition, the assessment items were designed to measure understanding of the takeaway points of each WEP, and those takeaways were more likely to be part of the whole class discussion than the small group work. Given the importance of opportunities for students to engage with mathematical ideas and explain their reasoning (*Common Core State Standards in Mathematics*, 2010) and past evidence for the effectiveness of working with peers in small groups, small group work – in addition to whole class discussion – is likely a useful instructional strategy.

Limitations and Future Direction

The current study helps advance our understanding of algebra learning and instruction, but much work remains to be done. First, the lack of random assignment of schools to condition and the limited success in matching schools in the treatment and control group on potentially important student demographic dimensions leaves open the possibility that the positive effects on student learning were not caused by treatment teachers using our approach. The lack of student-level demographic data further limited our ability to control for potential pre-existing differences between groups. Second, the teachers and their students were predominantly white and living in suburban communities, so the approach may need to be adapted to be successful in other contexts. For example, teachers may be more hesitant to compare multiple strategies with students they perceive to be less mathematically capable (Baxter et al., 2002). Third, linear equation solving is only one, albeit a core, algebra topic, and evidence is needed that a CDMS approach can promote learning of other algebra topics, such as solving systems of linear equations and polynomials and factoring. The current study provides evidence of promise of our CDMS approach to algebra instruction, and it is now important to test the effectiveness of our CDMS approach with a broader range of teachers and students, using random assignment to condition, and with a broader range of topics.

It is also important to continue to flesh out a theory of change for why CDMS improves teaching quality and student learning. For example, we hypothesize that students' and teachers' attitudes play a key role in understanding why comparison and discussion of multiple strategies can positively impact teachers' practice and student learning. Teachers' attitudes influence teachers' behaviors, including their teaching practices, and in turn impact student attitudes and learning (e.g., Chávez et al., 2015; Döhrmann et al., 2012; Midgley et al., 1989). Teachers' self-efficacy for helping students learn math is associated with their students having more positive attitudes about math (Midgley et al., 1989). Unfortunately, high school teachers often have unproductive views about why students struggle to learn math, attributing students' difficulties to characteristics of the students and their home environments and not to factors under the teachers' control, such as instructional quality (Jackson et al., 2017). These unproductive views are related to less productive strategies for remediating mathematical difficulties (e.g., repeated practice vs. teaching the content in a new way). Anecdotal evidence from informal teacher and student interviews suggests that using our CDMS approach provides materials and instructional routines that build teachers' self-efficacy for helping all students learn mathematics and positive beliefs about effective instructional practices (e.g., multiple strategies, comparison, flexibility, small group work, and discussion). For example, one teacher noted: "We did see a difference in their [the students'] engagement and their success. So, I'm a believer. Showing students multiple – because I think students are so used to, there's only one right way in math, and I think your program is super powerful to say, no, there's multiple ways." In the case of students, mathematically-capable students also need a productive set of attitudes, such as a productive disposition (i.e., "habitual inclination to see mathematics as sensible, useful, and worthwhile, coupled with a belief in diligence and one's own efficacy" (Kilpatrick et al., 2001, p. 116)). A flexibility mindset (meta-cognitive beliefs about procedural flexibility) may also be valuable (Authors, 2014a). For example, a student noted: "I liked how – I learned different ways to do it and there's different ways to solve the problem. I really liked that, how we went over both ways... Usually we just focus on one way and kinda just practice that and don't look at two ways to compare them." Future research needs to systematically assess teachers' and students' beliefs before and after participation in a CDMS approach. This evidence will help refine a theory of change for why CDMS improves student outcomes (e.g., that a CDMS curriculum and PD impacts teaching attitudes and practices, which in turn impacts student knowledge and attitudes).

In conclusion, the current study supports our hypothesis that productive learning of algebra is supported by reflection on multiple strategies through comparing them and discussing the rationale behind and relative merits of particular strategies. Our materials and PD promoted sustained comparison of multiple strategies, small group work and whole class discussions, and such practices were rare in control classrooms. They also promoted instruction that focused on making sense of procedures and developing procedural flexibility, in part through teachers asking conceptual and open-ended questions. Students in the treatment condition had higher total scores on the posttest than students in the control condition, and this difference was mainly due to treatment students having higher procedural flexibility. The findings are promising that encouraging teachers to regularly compare and discuss multiple strategies should increase students' algebra learning. Key contributions of the current study are highlighting the need to expand theories of algebra learning to include attention to procedural flexibility and illustrating an instructional theory and method to promote broader learning about algebra.

Data Availability Statement:
The data file as well as materials are available at
https://osf.io/89rj7/?view_only=2ba70793a8344204949767be1c03729b

References

Adelman, C. (2006). *The Toolbox Revisited: Paths to Degree Completion From High School Through College* (p. 223). Office of Vocational and Adult Education, Washington, DC.

Aleven, V. A. W. M. M., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, *26*(2), 147–179. https://doi.org/10.1207/s15516709cog2602_1

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning Through Case Comparisons: A Meta-Analytic Review. *Educational Psychologist*, *48*(2), 87–113. https://doi.org/10.1080/00461520.2013.775712

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, *70*, 181–214. https://doi.org/10.3102/00346543070002181

Australian Education Ministers. (2006). *Statements of learning for mathematics*. Curriculum Corporations.

Authors (2007). Details omitted for double-blind reviewing.

Authors (2008). Details omitted for double-blind reviewing.

Authors (2009a). Details omitted for double-blind reviewing.

Authors (2009b). Details omitted for double-blind reviewing.

Authors (2009c). Details omitted for double-blind reviewing.

Authors (2012). Details omitted for double-blind reviewing.

Authors (2013). Details omitted for double-blind reviewing.

Authors (2014a). Details omitted for double-blind reviewing.

Authors (2014b). Details omitted for double-blind reviewing.

Authors (2015a). Details omitted for double-blind reviewing.

Authors (2015b). Details omitted for double-blind reviewing.

Ball, D. L. (1993). With an Eye on the Mathematical Horizon: Dilemmas of Teaching Elementary School Mathematics. *The Elementary School Journal*, *93*, 373–397. https://doi.org/10.1086/461730

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An Alternative Reconceptualization of Procedural and Conceptual Knowledge. *Journal for Research in Mathematics Education*, *38*(2), 115–131. https://doi.org/10.2307/30034952

Baxter, J., Woodward, J., Voorhies, J., & Wong, J. (2002). We talk about it, but do they get it? *Learning Disabilities Research & Practice*, *17*(3), 173–185. https://doi.org/10.1111/1540-5826.00043

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study*. Boston College.

Blanton, M. L., & Kaput, J. J. (2005). Characterizing a classroom practice that promotes algebraic reasoning. *Journal for Research in Mathematics Education*, *36*(5), 412–446. https://doi.org/10.2307/30034944

Blöte, A. W., Van der Burg, E., & Klein, A. S. (2001). Students' flexibility in solving two-digit addition and subtraction problems: Instruction effects. *Journal of Educational Psychology*, *93*(3), 627–638. https://doi.org/10.1037/0022-0663.93.3.627

Blume, G. W., & Heckman, D. S. (1997). What do students know about algebra and functions? In P. A. Kenney & E. A. Silver (Eds.), *Results from the sixth mathematics assessment* (pp. 225–277). National Council of Teachers of Mathematics.

Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education*, 119–156.

Byrnes, J. P., & Wasik, B. A. (1991). Role of conceptual knowledge in mathematical procedural learning. *Developmental Psychology*, *27*(5), 777–786. https://doi.org/10.1037/0012-1649.27.5.777

Canobi, K. H. (2009). Concept–procedure interactions in children's addition and subtraction. *Journal of Experimental Child Psychology*, *102*(2), 131–149. https://doi.org/10.1016/j.jecp.2008.07.008

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, *26*(4), 499–531. https://doi.org/10.3102/00028312026004499

Carraher, D. W., Mara, Martinez., & Schliemann, A. D. (2008). Early algebra and mathematical generalization. *ZDM Mathematics Education*, 403–422. https://doi.org/10.1007/s11858-007-0067-7

Chávez, Ó., Tarr, J. E., Grouws, D. A., & Soria, V. M. (2015). Third-year high school mathematics curriculum: Effects of content organization and curriculum implementation. *International Journal of Science and Mathematics Education*, *13*(1), 97–120. https://doi.org/10.1007/s10763-013-9443-7

Chi, M. T. H. (2000). Self-explaining: The dual processes of generating inference and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (pp. 161–238).

Common Core State Standards in Mathematics. (2010). National Governors Association Center for Best Practices, Council of Chief State School Officers.

Döhrmann, M., Kaiser, G., & Blömeke, S. (2012). The conceptualisation of mathematics competencies in the international teacher education study TEDS-M. *ZDM*, *44*(3), 325–340. https://doi.org/10.1007/s11858-012-0432-z

Dowker, A. (1992). Computational estimation strategies of professional mathematicians. *Journal for Research in Mathematics Education*, *23*(1), 45–55. https://doi.org/10.2307/749163

Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction*, *22*, 206–214. https://doi.org/10.1016/j.learninstruc.2011.11.001

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, *95*, 393–405. https://doi.org/10.1037/0022-0663.95.2.393

Guo, J., & Pang, M. F. (2011). Learning a mathematical concept from comparing examples: The importance of variation and prior knowledge. *European Journal of Psychology of Education*, *26*(4), 495–525. https://doi.org/10.1007/s10212-011-0060-y

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K. C., Human, P., Murray, H., Olivier, A., & Wearne, D. (1996). Problem Solving as a Basis for Reform in Curriculum and Instruction: The Case of Mathematics. *Educational Researcher*, *25*(4), 12–21. https://doi.org/10.3102/0013189X025004012

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., Chiu, A. M. Y., Wearne, D., Smith, M., & Kersting, N. (2003). Teaching mathematics in seven countries: Results from the TIMSS 1999 video study (NCES 2003-013). *US Department of Education. Washington, DC: National Center for Education Statistics*.

Hodds, M., Alcock, L., & Inglis, M. (2014). Self-Explanation Training Improves Proof Comprehension. *Journal for Research in Mathematics Education*, *45*(1), 62–101. https://doi.org/10.5951/jresematheduc.45.1.0062

Jackson, K., Gibbons, L., & Sharpe, C. J. (2017). Teachers' views of students' mathematical capabilities: Challenges and possibilities for ambitious reform. *Teachers College Record*, *119*(7), 1–43.

Johnson, D. W., & Johnson, R. T. (1994). *Learning together and alone: Cooperative, competitive and individualistic learning* (4th ed.). Allyn and Bacon.

Kaput, J. J. (1998). Representations, inscriptions, descriptions and learning: A kaleidoscope of windows. *The Journal of Mathematical Behavior*, *17*(2), 265–281. https://doi.org/10.1016/S0364-0213(99)80062-7

Kieran, C. (1992). The learning and teaching of school algebra. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 390–419). Simon & Schuster.

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

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does Understanding the Equal Sign Matter? Evidence from Solving Equations. *Journal for Research in Mathematics Education*, *37*(4), 297–312. https://doi.org/10.2307/30034852

Kultusministerkonferenz. (2004). *Bildungsstandards im Fach Mathematik für den Primarbereich [Educational Standards in Mathematics for Primary Schools]*. Luchterhand.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal*, *27*, 29–63. https://doi.org/10.3102/00028312027001029

Lindquist, M. M. (1989). *Results from the fourth mathematics assessment of the National Assessment of Educational Progress*. National Council of Teachers of Mathematics.

Litke, E. (2020). The Nature and Quality of Algebra Instruction: Using a Content-Focused Observation Tool as a Lens for Understanding and Improving Instructional Practice. *Cognition and Instruction*, *38*(1), 57–86. https://doi.org/10.1080/07370008.2019.1616740

Lloyd, G. M., Herbel-Eisenmann, B. A., & Star, J. R. (2011). *Developing essential understanding of expressions, equations, and functions for teaching mathematics in grades 6-8*. National Council of Teachers of Mathematics.

Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology*, *81*(2), 247. https://doi.org/10.1037/0022-0663.81.2.247

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General*, *131*(1), 5–15. https://doi.org/10.1037/0096-3445.131.1.5

National Center for Education Statistics. (2020a). *NAEP report card: 2019 NAEP mathematics assessment*. The Nation's Report Card. https://www.nationsreportcard.gov/highlights/mathematics/2019/

National Center for Education Statistics. (2020b). *NAEP Questions Tool*. National Assessment of Educational Progress. https://nces.ed.gov/NationsReportCard/nqt/

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. National Council of Teachers of Mathematics, Inc.

National Mathematics Advisory Panel. (2008). *Foundations of Success: The Final Report of the National Mathematics Advisory Panel*. U.S. Department of Education.

Newton, K. J., Star, J. R., & Lynch, K. (2010). Understanding the development of flexibility in struggling algebra students. *Mathematical Thinking and Learning*, *12*(4), 282–305. https://doi.org/10.1080/10986065.2010.482150

OECD. (2013). Chapter 2. A Profile of Student Performance in Mathematics. In *PISA 2012 Results: What Students Know and Can Do (Volume I): Student Performance in Mathematics, Reading and Science*. OECD Publishing. https://doi.org/10.1787/9789264201118-5-en.

Richland, L. E., Holyoak, K. J., & Stigler, J. W. (2004). Analogy use in eighth-grade mathematics classrooms. *Cognition and Instruction*, *22*, 37–60. https://doi.org/10.1207/s1532690Xci2201_2

Richland, L. E., Stigler, J. W., & Holyoak, K. J. (2012). Teaching the Conceptual Structure of Mathematics. *Educational Psychologist*, *47*(3), 189–203. https://doi.org/10.1080/00461520.2012.667065

Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. *Science*, *316*(5828), 1128–1129. https://doi.org/10.1126/science.1142103

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

Rittle-Johnson, B., Saylor, M., & Swygert, K. E. (2008). Learning from explaining: Does it matter if mom is listening? *Journal of Experimental Child Psychology*, *100*(3), 215–224. https://doi.org/10.1016/j.jecp.2007.10.002

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. https://doi.org/10.1037/0022-0663.93.2.346

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology*, *99*(3), 561–574. https://doi.org/10.1037/0022-0663.99.3.561

Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology*, *101*, 529–544. https://doi.org/10.1037/a0014224

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2009). The importance of prior knowledge when comparing examples: Influences on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology*, *101*(4), 836–852. https://doi.org/10.1037/a0016026

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2012). Developing procedural flexibility: Are novices prepared to learn from comparing procedures? *British Journal of Educational Psychology*, *82*(3), 436–455. https://doi.org/10.1111/j.2044-8279.2011.02037.x

Schmidt, W. H., McKnight, C. C., Cogan, L. S., Jakwerth, P. M., & Houang, R. T. (1999). *Facing the consequences: Using TIMSS for a closer look at US mathematics and science education*. Kluwer.

Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations between conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology*, Advance online publication. https://doi.org/10.1037/a0024997

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*, 475–522. https://doi.org/10.1207/s1532690xci1604_4

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. Oxford University Press.

Silver, E. A., Ghousseini, H., Gosen, D., Charalambous, C., & Strawhun, B. (2005). Moving from rhetoric to praxis: Issues faced by teachers in having students consider multiple solutions for problems in the mathematics classroom. *Journal of Mathematical Behavior*, *24*, 287–301. https://doi.org/10.1016/j.jmathb.2005.09.009

Singapore Ministry of Education. (2006). *Secondary mathematics syllabuses*.

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education*, *36*, 404–411. https://doi.org/10.2307/30034943

Star, J. R., Caronongan, P., Foegen, A. M., Furgeson, J., Keating, B., Larson, M. R., Lyskawa, J., McCallum, W. G., Porath, J., & Zbiek, R. M. (2015). *Teaching strategies for improving algebra knowledge in middle and high school students*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U. S. Department of Education.

Star, J. R., & Newton, K. J. (2009). The nature and development of experts' strategy flexibility for solving equations. *ZDM-International Journal on Mathematics Education, 41*, 557–567. https://doi.org/10.1007/s11858-009-0185-5

Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology*, *102*, 408–426. https://doi.org/10.1016/j.jecp.2008.11.004

Star, J. R., & Seifert, C. (2006). The development of flexibility in equation solving. *Contemporary Educational Psychology*, *31*, 280–300. https://doi.org/10.1016/j.cedpsych.2005.08.001

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating Productive Mathematical Discussions: Five Practices for Helping Teachers Move Beyond Show and Tell. *Mathematical Thinking and Learning*, *10*(4), 313–340. https://doi.org/10.1080/10986060802229675

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455–488. https://doi.org/10.3102/00028312033002455

Treffers, A. (1991). *Realistic mathematics education in the Netherlands 1980-1990*.

Verschaffel, L., Luwel, K., Torbeyns, J., & Van Dooren, W. (2009). Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education. *European Journal of Psychology of Education*, *24*(3), 335. https://doi.org/10.1007/BF03174765

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education, 22*(5), 366-389.

Webb, N. M., Franke, M. L., Ing, M., Wong, J., Fernandez, C. H., Shin, N., & Turrou, A. C. (2014). Engaging with others' mathematical ideas: Interrelationships among student participation, teachers' instructional practices, and learning. *International Journal of Educational Research*, *63*, 79–93. https://doi.org/10.1016/j.ijer.2013.02.001

Woodward, J., Beckmann, S., Driscoll, M., Franke, M. L., Herzig, P., Jitendra, A. K., Koedinger, K. R., & Ogbuehi, P. (2012). *Improving mathematical problem solving in grades 4 through 8: A practice guide*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U. S. Department of Education.

Table 1

Sample Items on Linear Equation Solving Assessment

<table>
<tr><th>Knowledge Type</th><th>Sample Items</th></tr>
<tr><td>I. Procedural Flexibility
(n = 6)</td><td>α = .53</td></tr>
<tr><td>a.</td><td>Below is the beginning of Gabriella’s, Jamal’s, and Nadia’s work in solving the equation $x + 7 - 3 = 12 - 2x$. To start solving this problem, which way(s) may be used?
a. Gabriella’s way
b. Jamal’s way
c. Nadia’s way
d. Jamal’s and Nadia’s ways
e. Gabriella’s, Jamal’s, and Nadia’s ways
<table>
<tr><td>Gabriella’s way:

Subtract 3 from 7:
$x+4=12-2x$</td><td>Jamal’s way:

Add 2x to both sides:
$3x+7-3=12$</td><td>Nadia’s way:

Subtract (7 – 3) from both sides:
$x=8-2x$</td></tr>
</table>
</td></tr>
<tr><td>b.</td><td>On a timed test, which would be the BEST way to solve the equation below for x given the values for y listed in the table? (Circle the letter for the best way).
$4x + 18 = y$
<table>
<tr><th>y</th></tr>
<tr><td>26</td></tr>
<tr><td>30</td></tr>
<tr><td>22</td></tr>
<tr><td>6</td></tr>
<tr><td>14</td></tr>
</table>
<table>
<tr><td>a. Gabriella’s way:

$4x + 18 = 26$
y = 26: $4x = 26 - 18$
…

$4x + 18 = 30$
y = 30: $4x = 30 - 18$
…

y = 22: …</td><td>b. Jamal’s way:

$4x + 18 = 26$
y = 26: $\frac{4x + 18}{4} = \frac{26}{4}$
…

$4x + 18 = 30$
y = 30: $\frac{4x + 18}{4} = \frac{30}{4}$
…

y = 22: …</td><td>c. Nadia’s way:

$4x = y - 18$
$x = \frac{y - 18}{4}$

y = 26:
$x = \frac{26 - 18}{4}$

y = 30:
$x = \frac{30 - 18}{4}$

y = 22: …</td></tr>
</table>
</td></tr>
</table>

Knowledge Type	Sample Items
II. Procedural Knowledge (n = 5)	α = .60
a.	Solve the equation below for *x*. Circle the letter for your answer. $3(2x + 3x - 4) + 5(2x + 3x - 4) = 48$ a. $\frac{7}{8}$ b. -1 c. $\frac{2}{5}$ d. 2
b.	Solve the equation below for *y*. $5(y - 2) = -3(y - 2) + 4$
III. Conceptual Knowledge (n = 5)	α = .45
a.	Which of the following is a like term to (could be combined with) **7(*j* + 4)**? a. $7(j + 10)$ b. $2(j + 4)$ c. $7(p + 4)$ d. Both a and b e. All of the above
b.	Look at this pair of equations. **Without solving the equations**, decide if these equations are equivalent (have the same answer). $6(x + 3) = 60$ $x + 3 = 10$ a. YES (same answer) b. NO (different answer) c. CAN’T TELL without doing the math d. CAN’T TELL because I need more information

Note. Cronbach’s alphas at posttest.

Table 2

Percent of Lessons with Each Feature of our General Fidelity Coding, by Condition

Code	% of Treatment	% of Control
Exposed students to multiple strategies	100	4
Multiple strategies were presented side-by-side	100	0
Multiple strategies were compared for at least a 1.5-minute continuous block	97	0
Engaged in partner/small group work focused on math content for at least a 1-minute continuous block	90	46
Had a whole class discussion for at least a 1.5-minute continuous block	83	13

Table 3

Average Rating of Implementation Quality When Using CDMS Materials on a Scale from 1-4

Code	Description	Description of mean rating	Mean	SD
Making sense of procedures	The extent that the teacher's explanations and/or questions were intended to push students toward making sense of strategies.	Teacher asks questions or provides explanations focused on sense-making occasionally.	2.56	1.03
Supporting procedural flexibility	The extent to which teachers presented procedures so students had the opportunity to develop flexibility, particularly focusing on multiple strategies and considering which strategies to use on certain problems.	A focus on procedural flexibility clearly happens in explicit ways.	3.09	0.90
Teacher questioning	The extent that the teacher created an opportunity for students to engage in deep and sustained mathematical thinking with their prompts.	Dominated by questions (e.g., why) where students are expected to provide answers that are longer than a word but where generally there is a right and a wrong answer. Some more open-ended questions may be used.	3.09	0.52
Student responses	The extent that the classroom environment created by the teacher was one where students felt comfortable expressing themselves and that a variety of students did so – that students were inspired to contribute in response to questions from the teacher.	The nature of students' responses is a mix of short (one word or a short sentence) and long – where a long response is when a single student holds the floor for about 15 seconds or more.	2.88	0.72
Opportunities for interaction	Degree to which the teacher created a classroom environment where students began engaging in mathematical talk with each other and not only with the teacher.	The teacher's attempts to stimulate student interaction through her questions occur infrequently and may include tactics such as asking multiple students the same question.	1.56	0.73

Note: See Supplemental Materials for more detailed coding scheme.

Table 4

Multilevel Regression Models Results for Predicting Students' Posttest Scores

	Overall				Conceptual				Procedural				Flexibility			
	B	SE	*t*	*p*	*B*	SE	*t*	*p*	*B*	SE	*t*	*p*	*B*	SE	*t*	*p*
Intercept	7.32	0.35	20.98	.000	2.25	0.13	16.90	.000	2.70	0.24	11.10	.031	2.41	0.19	12.55	.000
Condition	1.29	0.64	2.02	.050	0.38	0.24	1.54	.130	0.11	0.40	0.28	.811	0.79	0.35	2.25	.030
Pretest score	0.61	0.04	16.33	.000	0.41	0.04	11.78	.000	0.43	0.03	13.13	.000	0.37	0.04	8.51	.000
%FRPL	0.05	0.03	1.49	.143	0.02	0.01	1.72	.094	0.02	0.02	1.28	.295	0.02	0.02	1.36	.181
%African American	0.12	0.07	1.79	.081	0.06	0.03	2.36	.023	0.03	0.04	0.89	.445	0.02	0.04	0.65	.518
%Hispanic	-0.09	0.04	-1.98	.055	-0.04	0.02	-2.42	.020	-0.04	0.02	-1.89	.165	-0.03	0.02	-1.17	.248

Note. FRPL stands for percent of students at school who received free or reduced price lunch.

Figure 1

Sample Worked Example Pair

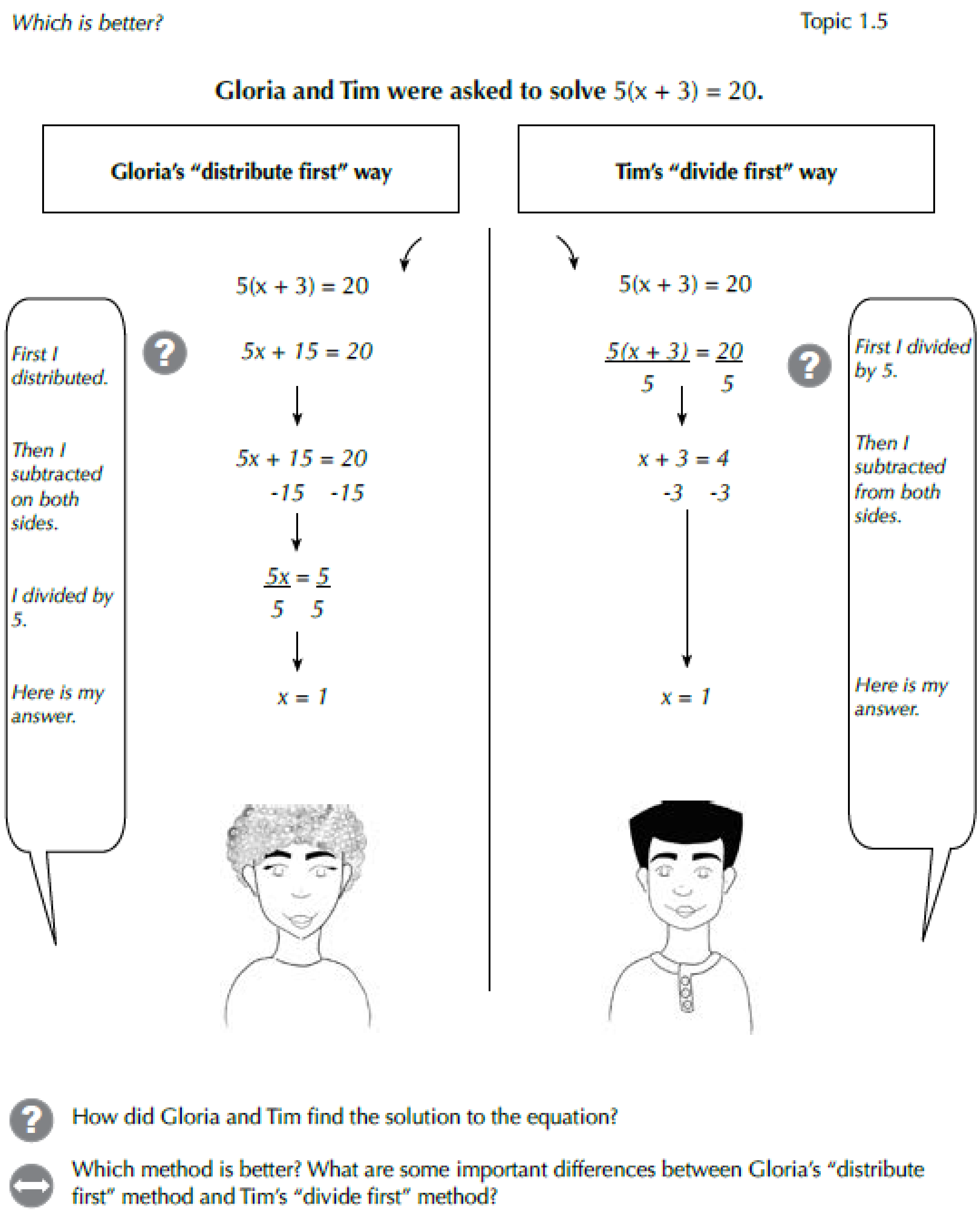


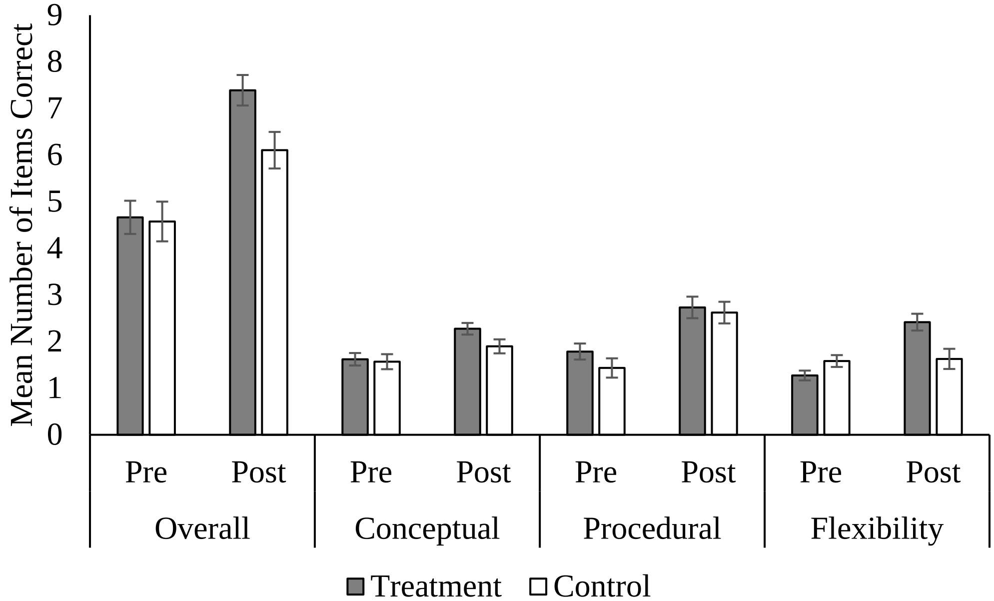
Figure 2

CDMS Implementation Model

Compare (8 minutes)	**Prepare to Compare** ➢ What is the problem asking? ➢ What is happening in the first method? ➢ What is happening in the second method?
	Make Comparisons ➢ What are the similarities and differences between the two methods? ○ Which method is better? ○ Which method is correct? ○ Why do both methods work? ○ How do the problems differ?
Discuss (12minutes)	**Prepare to Discuss (think, pair)** ➢ How does this comparison help you understand this problem? ➢ How might you apply these methods to a similar problem?
	Discuss Connections (share) ➢ What ideas would you like to share with the class?
	Identify the Big Idea ➢ Can you summarize the Big Idea in your own words?

Figure 3

Mean Number of Pretest and Posttest Items Correct, by Condition



Note. Estimated marginal means of total items (out of 16), conceptual items (out of 5), procedural items (out of 5), and flexibility items (out of 6) correct on unit pretest and posttest by condition.

List of Figures

Supplemental Materials

Appendix A

General Fidelity Coding Scheme

1a. Were students exposed to multiple strategies?

Check 'Yes' if:

More than one way for solving a given problem was present during a single class period. This includes instances where the text presents multiple strategies and the teacher describes what is in the text.

Check 'No' if:

More than one way for solving a given problem was not present during a single class period, or if an alternative strategy was only briefly (less than 10 seconds) mentioned. Simply mentioning that there is another strategy, without describing it in some detail (for at least 10 seconds), does not count. Note that if the response to this question is no, skip to question 2.

1b. Were the multiple strategies presented side-by-side?

Check 'Yes' if:

The strategies are visually shown to students side-by-side on a worksheet, board, or overhead, so that both strategies are visible to students at the same time.

Check 'No' if:

The multiple strategies are not shown side by side and are not visible to students at the same time.

1c. Did the teacher or students compare the multiple strategies for at least a 1.5-minute continuous block?

Check 'Yes' if:

Regardless of whether the strategies are presented side-by-side, the teacher or student(s) engages the class in thinking about how the two strategies are similar or different for at least 1.5 minute continuous block. Comparison of strategies might include mention of when a strategy is useful or not useful as compared to another strategy, when a strategy might result in a more or less errors as compared to another strategy, when a strategy might be more or less applicable as compared to another strategy, and/or why both strategies work. The comparison can be implied (e.g., talking about one solution being better and why without explicitly talking about the other solution).

Check 'No' if:

Multiple strategies are considered, but not compared, even if the strategies are side-by-side. Or strategies are compared, but for less than 1.5 minutes at a time. Note that several short segments of comparison, added together, do not count. It must be a sustained comparison for at least 1.5 minutes.

2. Did <u>all</u> students engage in partner or small group work focused on math content for at least a 1-minute continuous block?

Check 'Yes' if:

Students work together on math content in a small group of 2 or more students for at least a 1-minute continuous block. Do not count time during which the **teacher** engages in off-task talk that disrupts small group work.

Check "No" if:

Small group work was optional, so only a subset of students worked in small groups, small group work focused on math content was brief, lasting less than a 1-minute continuous block, or small group work was not present.

3. Was there a whole class discussion for at least a 1.5-minute continuous block?

Setting is the *whole class* (not a small group)

Check 'Yes' if:

For at least a 1.5-minute continuous block, *as least one* of these things was happening (a) teacher is asking <u>conceptual or open-ended questions</u> and more than one student is responding to the questions (multiple students do not have to answer the same question) and/or (b) teacher is redirecting conversation by following up on a student's response to ask another student to respond to the same question or to the previous student's idea. Teachers asking questions and explaining and elaborating on what students said is okay.

Conceptual questions (often why questions) include why an answer was correct or why/when a particular solution method might have been a good choice (e.g. "Why is her method better?; What made Alex's answer correct?"),

Open-ended questions do not have a predetermined answer and students can have different takeaway points from the same question, such as "What's another perspective?", "Why would you use a different method if the ordered pair changed to (-3000, 52)?", "Can you generate a new problem where Riley's method could not be used?", "Will Tim's method always work and why or why not?", "What is the big idea or takeaway?"

Reminder: If teacher is doing part b – redirecting conversation – then the questions may be focused on other things, such as how questions (e.g., "how did you get that?").

Start time of discussion: Teacher asks a question of substance (open-ended question, conceptual question, when, why, etc.) with intent that students will respond to that question.

End time of discussion: Teacher talks for 1 minute or more or activity of the class shifts, such as addressing a new problem, moving to individual work (e.g., homework, write big idea on sheet with no class discussion after). To determine exact end time, look back for when students stopped answering questions.

Do not count as questions: Level 1 questions or rhetorical questions (e.g., Everyone understand? Okay?).

Note that it is possible for lesson time to count towards fulfilling two different codes. Questions such as, “Why is her method better?” and “Why does this method work?” qualify as both comparison and discussion.

Check ‘No’ if:

Any discussion only happens in a small group setting, not a whole class setting.

Any whole class discussion is brief, lasting for less than a 1.5-minute continuous block.

The discussion is not about mathematical content (e.g., about logistics).

The teacher asks questions, but does not wait for students to respond to the questions, sometimes answering the questions herself.

The segment involves minimal student participation, such as students only giving one word answers (e.g., yes, no, Alex, distributive property) or stating the solution or formula (e.g., $y = 2$, $7a + 5$, $y = mx + b$). The teacher’s ideas drive the conversation.

The segment involved only one student responding to the teacher’s questions.

Appendix B

Treatment Implementation Quality Coding Scheme

MAKING SENSE OF PROCEDURES

This code is intended to capture the extent that the teacher's explanations and/or questions are intended to push students toward making sense of procedures and strategies in the WEP portion of the lesson and refers to deliberate actions that the teacher takes. Some WEPs are explicitly designed with questions and dialogue focused on making sense of procedures, while others are not. Some of what the teacher may focus on – in the form of questions and/or explanations – that signal her interest in students' sense-making of procedures are the following:

- The WHY that supports individual steps in a procedure (e.g., WHY you plug in x = 0 into a linear equation when finding the y-intercept)
- The WHY that explains the solution generated by a procedure (e.g., when the ordered pair (x, y) is a solution to a system of linear equations, this means that (x, y) is the point of intersection of two lines and/or results in a true statement when plugged into both equations)
- The purpose/mathematical goal of a procedure (e.g. using quadratic formula allows us to find the roots of a parabola)
- The mathematical properties underlying a procedure (e.g., how FOIL is really the distributive property, how y = 2 is a horizontal line because all of its point have the form (x, 2); that shaded points in graphs of inequalities indicate values that make the inequality true)
- The WHY indicating the reasons that a procedure holds (e.g. when you multiply exponents with a common base, you can add them because multiplication works as repeated addition)

Also, note that this code is intended to capture teachers' efforts to make sense of procedures in the *whole class portions* of the class, *not* in partner or group work.

1– Little or no focus	2 – Low – incidental focus	3 – Medium – moderate focus	4 – High – major and sustained focus
Includes little or no indication that the teacher is interested in having students make sense of procedures. If there are sense-making questions or explanations in the WEP, the teacher does not go beyond a rare brief comment.	While making sense of procedures, in the form of teacher questions and explanations, occurs *occasionally* or *incidentally*, it is not sustained or an explicit focus of the instruction. The teacher asks questions or provides explanations focused on sense-making – either those in the WEP or supplements. Even if there are multiple instances of such questions and explanations, these are relatively infrequent, short in duration, and done in passing.	Making sense of procedures clearly happens in explicit ways. This focus is neither incidental – occurring occasionally or in passing – nor is it a sustained major focus of the lesson. Rather, sense-making occurs for one sustained time or for several times, including questions and explanations that are part of the WEP but perhaps supplements as well.	Making sense of procedures is a prominent, explicit, and *major focus* of the WEP portion of the class. The teacher not only utilizes questions and explanations included in the WEP in pursuit of this focus *but also supplements with additional explanations and questions pushing students to make sense of procedures.*

SUPPORTING PROCEDURAL FLEXIBILITY

This code is intended to capture the extent to which teachers present procedures and strategies such that students had the opportunity to develop procedural flexibility, particularly focusing on multiple strategies and working with students to consider which strategies to use on certain problems, and this code focuses on the actions that the teacher takes in support of procedural flexibility. Note that most WEPs contain some built-in support for procedural flexibility, since multiple strategies are always presented. And some WEPs are explicitly focused on procedural flexibility, particularly "Which is better?" WEPs. Also, note that this code is intended to capture teachers' efforts to support procedural flexibility in the *whole class portions* of the class, *not* in partner or group work.

In supporting procedural flexibility, the teacher may:

- Discuss multiple strategies for approaching the same problem, perhaps with a focus on when a particular strategy may be especially beneficial or efficient to use
- Attend to applicability conditions of a procedure (e.g. by noting when it can or can't be used or what problem conditions led to the choice of a given procedure)
- Attend to the key conditions of steps within a procedure to be able to understand its usefulness/efficiency in specific situations as opposed to other situations
- Use a heuristic or identify a problem type for evaluating when a procedure is useful or efficient (e.g., when we see problems that look like <problem feature> it means this strategy might be a good idea)

1 – Little or no focus	2 – Low – incidental focus	3 – Medium – moderate focus	4 – High – major and sustained focus
Includes little or no indication that the teacher is interested in having students develop procedural flexibility. The teacher does not go beyond a rare brief comment related to flexibility in these strategies.	Procedural flexibility is an incidental or occasional focus. This may occur when the WEP is explicitly focused on flexibility but the teacher does not dive into or dwell on flexibility. It may also occur when the WEP is not focused on flexibility but the teacher occasionally asks questions or makes short explanations related to flexibility. Even if there are multiple instances of such questions and explanations, these are relatively infrequent, short in duration, and done in passing.	A focus on procedural flexibility clearly happens in explicit ways during the WEP implementation. This focus is neither incidental – occurring occasionally or in passing – nor is it a sustained major focus of the lesson. Rather, emphasis on flexibility occurs for one sustained time or in several times, including questions and explanations that are part of the WEP but perhaps supplements as well.	Procedural flexibility is a prominent, explicit, and *major focus* of the WEP portion of the class. The teacher not only utilizes questions and explanations included in the WEP in pursuit of this focus but *also supplements with additional explanations and questions pushing students to be flexible.*

TEACHER QUESTIONING

This code is intended to capture the extent that the teacher (via questioning) creates an opportunity for students to engage in deep and sustained mathematical thinking. (These types of opportunities for deep thinking are presumed to occur as a result of the types of questions that teachers ask.) The coding levels refer to the kinds of teacher questions that are most salient or instrumental in the mathematical work of the lesson. Questions asked that do not play a role in the mathematical work of the class are not considered (e.g., logistical questions). We also note that we consider as questions only those statements from the teacher that are asked with the interest of being answered – meaning that rhetorical questions (e.g., "Alright?") or questions asked without any pause or attention to the possibility that students might answer (e.g., "Any questions?" without a pause for anyone to answer) are not counted as questions. We consider the following framework for questions.

Type 1 questions are yes/no questions or, more generally, questions that can be (and may indeed be) answered with a single word or number. Type 2 questions can generally be answered within a sentence and typically have a clear right or wrong answer.

Type 3 questions are open-ended questions, often require longer answers, and generally do not have a pre-established or right/wrong answer. Also, note that this code is intended to capture teacher questioning in the *whole class portions* of the class, *not* in partner or group work.

1 – Little or no questioning	2 – Mostly Type 1 questions	3 – Critical mass of Type 2 questions	4 – Critical mass of Type 3 questions
A teacher is not asking students questions but instead is generally doing the talking herself. When questions included as part of the WEP are asked, they are asked rhetorically such that there is no clear expectation that students will answer and/or the teacher answers the questions herself.	Lesson is dominated by Type 1 questions. The teacher poses questions that students are intended to answer, but the answers provided or required are generally short (e.g., yes/no, or numbers). There may be some higher-level questions, such as those included as part of the WEP. But on the whole, the majority of the lesson revolves around the teacher's use of Type 1 questions.	Lesson is dominated by Type 2 questions – where students are expected to provide answers that are longer than a word but where generally there is a right and a wrong answer. Some Type 3 questions may be used – particularly those included as part of the WEP. But the use of Type 2 questions is a substantial component of the lesson, including the teacher supplementing the provided questions with additional questions of Type 2.	The lesson contains a significant number of Type 3 questions, where students are asked to elaborate, to speak for more than one sentence, and to make interpretations or judgments. There may be lower level questions used by the teacher, but the presence of (and time spent asking and answering) Type 3 questions is a substantial part of the lesson. This usually requires teachers asking supplemental Type 3 questions not in the WEP materials.

STUDENT RESPONSES

This code is intended to capture the extent that the classroom environment created by the teacher is one where students feel comfortable expressing themselves and *that a variety of students do so* – that students are inspired to contribute in response to mathematical questions from the teacher. Because of poor student audio, it is generally not possible to hear what students are saying. So in this code, it may often be necessary to infer the nature of students' responses based on how teachers respond to the students. Also note that we are only interested in students' responses to mathematical questions. The code focuses on the characterization of students' responses to teachers' questions during the lesson, including how many students are responding to questions, the length of each student's turn while talking, and the content of students' contributions (when it is possible to hear them). Also, note that this code is intended to capture student responses in the *whole class portions* of the class, *not* in partner or group work.

1 – Little or no individual responses	2 – Regular short individual responses	3 – Mix of short and long individual responses	4 – Substantial and elaborated responses from many students
Almost entirely focused on teacher talk. Students' responses are limited to 'choral' (group) responses to teachers' questions or occasional individual (e.g., called upon by name or hands raised) responses to Type 1 (yes/no) question. The total number of students in the class who are participating by offering individual (called by name) responses is small.	Students respond to the teacher's questions regularly throughout the WEP portion of the lesson. But the nature of students' responses is mostly in the form of single words or short sentences. A variety of students in the class are offering individual responses – e.g., many students in the class are called upon to participate.	Students respond to the teacher's questions regularly throughout the lesson. The nature of students' responses is a mix of short (one word or a short sentence) and long – where a long response is when a single student holds the floor for about 15 seconds or more. The lesson may include a few instances where one or more students offer longer responses, yet only a small number (one or two) students offer these longer responses. Yet a relatively large number of students are called upon to participate generally (attend to whether this last sentence should be kept in code)	Lesson is characterized by several students taking relatively long speaking turns in response to teachers' questions. Students are regularly responding to teachers' questions during the lesson, and there may be some other forms of responses (e.g., short or one-word responses). But in general, a noteworthy feature of the lesson is that students are talking in long turns and the teacher is asking questions and listening a lot to students' contributions.

OPPORTUNITIES FOR STUDENT INTERACTION

The interaction code is intended to assess the degree to which the teacher creates a classroom environment where students begin engaging in mathematical talk with each other and not only with the teacher. By virtue of the ways that she responds to students' utterances, the teacher not only asks good questions (captured in the teacher questioning code) and the students not only feel comfortable responding (captured in the student responses code), but the teacher also encourages students to listen to, interact with, and respond to each other. Among the strategies that the teacher could use to push students in this direction are deflecting a question directed at the teacher and posing it back to a student, asking a student to rephrase what another student has said, and asking a student whether she disagrees with another student and why. Because we usually cannot hear students' utterances, this code does not consider whether the teacher's encouragement efforts in this direction are fruitful. Also, note that this code is intended to capture student interaction in the *whole class portions* of the class, *not* in partner or group work.			
1 – Little or no teacher attempts to encourage interaction	2 – Low - Occasional and/or infrequent teacher attempts to encourage interaction	3 – Medium - Moderate teacher attempts to encourage interaction	4 – High- Major and sustained teacher attempts to encourage interaction
The teacher does not attempt (in her use of questioning) to encourage student interaction or her limited attempts are not successful.	The teacher's attempts to stimulate student interaction through her questions occur infrequently and may include tactics such as asking multiple students the same question.	Teacher attempts to stimulate student interaction through tactics such as rephrasing student contributions in order to direct them to other students, or asking other students to rephrase a student's work, clearly happens in explicit ways. This focus is neither incidental – occurring occasionally or in passing – nor is it a sustained major focus of the lesson.	Teacher attempts to stimulate student interaction through tactics such as rephrasing student contributions in order to direct them to other students, or asking other students to rephrase a student's work is a prominent, explicit, and *major focus.*