RUNNING HEAD: Power of comparison in mathematics instruction

The Power of Comparison in Mathematics Instruction: Experimental Evidence from Classrooms

Bethany Rittle-Johnson

Jon Star

Kelley Durkin

Citation: Rittle-Johnson, B., Star, J. R., & Durkin, K. (2017). The power of comparison in mathematics instruction: Experimental evidence from classrooms. In D. C. Geary, D. B. Berch, & K. M. Koepke (Eds.), *Mathematical Cognition and Learning* (Vol. 3). (pp. 273-296). Waltham, MA: Elsevier.

Acknowledgement

Bethany Rittle-Johnson, Department of Psychology and Human Development, Vanderbilt University. Jon R. Star, Graduate School of Education, Harvard University. Kelley Durkin, Peabody Research Institute, Vanderbilt University.

Much of the research reported in this article was supported by grants from the National Science Foundation (DRL0814571) and the Institute of Education Sciences (R305H050179); the ideas in this paper are those of the authors and do not represent official positions of NSF or IES.

Correspondence concerning this article should be addressed to Bethany Rittle-Johnson, 230 Appleton Place, Peabody #0552, Nashville, TN 37203, 615-322-8301 (voice), b.rittle-johnson@vanderbilt.edu

Abstract

Comparison is a fundamental cognitive process that supports learning in a variety of domains. To leverage comparison in mathematics instruction, evidence-based guidelines are needed for how to use comparison effectively. In this chapter, we review our classroom-based research on using comparison to help students learn mathematics. In five short-term experimental, classroom-based studies, we evaluated two types of comparison for supporting the acquisition of mathematics knowledge and tested whether prior knowledge moderated their effectiveness. Comparing different solution methods for solving the same problem was particularly effective for supporting procedural flexibility across students and for supporting conceptual and procedural knowledge among students with some prior knowledge of one of the methods. We next developed a supplemental Algebra 1 curriculum to foster comparison and evaluated its effectiveness in a randomized-control trial. Teachers used our supplemental materials much less often than expected, and student learning was not greater in classrooms that had been assigned to use our materials. Students' procedural knowledge was positively related to greater implementation of the intervention, suggesting the approach has promise when used sufficiently often. This study suggests that teachers may need additional support in deciding *what* to compare and *when* to use comparison.

KEYWORDS: Mathematics learning; mathematics instruction; learning processes; comparison; classroom-based research; algebra

The Power of Comparison in Mathematics Instruction: Experimental Evidence from Classrooms

We often learn through comparison. For example, we compare new words, objects and ideas to ones we already know, and these comparisons help us recognize what features are important and merit more attention. Indeed, comparison aids learning across a broad array of topics, ranging from babies learning the distinction between dogs and cats (Oakes & Ribar, 2005), to preschoolers learning new words (e.g., Namy & Gentner, 2002), to business-school students learning contract negotiation skills (Gentner, Loewenstein, & Thompson, 2003). A recent meta-analysis confirmed that comparison promotes learning across a range of domains (Alfieri, Nokes-Malach, & Schunn, 2013). As Goldstone and colleagues noted: "Comparison is one of the most integral components of human thought…. Furthermore, research has demonstrated that the simple act of comparing two things can produce important changes in our knowledge" (2010, p. 103).

In this chapter, we focus on using comparison to support mathematics learning. Comparison is integral to best practices in mathematics education. Having students share and compare solution methods for solving a particular problem (e.g., discuss the similarities and differences in the methods) lies at the core of reform pedagogy in many countries throughout the world (Australian Education Ministers, 2006; Brophy, 1999; Kultusministerkonferenz, 2004; NCTM, 2014; Singapore Ministry of Education, 2006; Treffers, 1991), including the Common Core State Standards in Mathematics in the U.S. (2010). Its inclusion as a best practice was based on observational research that expert teachers in the U.S., as well as teachers from high-performing countries, have students compare multiple methods for solving problems during mathematics instruction (Ball, 1993; Lampert, 1990; Richland, Zur, & Holyoak, 2007).

Based on this convergence in the cognitive science and mathematics education literatures, we identified comparison as a promising instructional practice for improving mathematics learning. Our goal was to evaluate how comparison supports learning of school mathematics within a classroom setting. In this chapter, we present results from our classroom-based research on using comparison to improve mathematics learning, supplementing with recent research by others. We begin with short-term, researcher-led classroom research. Then, we describe a year-long study on helping teachers use comparison throughout the Algebra I curriculum.

Short-Term, Researcher-Led Classroom Research

In our initial research on the effectiveness of comparison, we redesigned 2-3 middle-school math lessons on a particular topic in several different ways and implemented these lessons during students' mathematics classes. This approach ensured that students were accountable for learning the material and that the material was readily usable in a typical classroom setting. We evaluated two types of comparison: comparing multiple methods for solving the same problem (*comparing methods*) and comparing different problems solved with the same method (*comparing problems*).

Typically, students in the control condition studied the same material sequentially, without comparison. This allowed us to isolate the effectiveness of comparison. Students worked with partners within the same classroom, and each pair of students was randomly assigned to a comparison or sequential condition. By randomly assigning pairs of students within the same classroom to a condition, we avoided confounding classroom effects with condition effects.

Instructional Materials

Before reviewing individual studies, we describe the design of the instructional materials, both for the comparison and control conditions. Worked examples along with prompts to explain the examples were the core of our instructional materials across experimental and control conditions. Worked examples present solution methods step-by-step and are a very effective way to help novices learn new procedures and related concepts (Atkinson, Derry, Renkl, & Wortham, 2000; Sweller & Cooper, 1985). They are commonly used in textbooks, so they are also familiar to students. Worked examples are also an effective way to introduce students to alternative, more efficient methods (Star & Rittle-Johnson, 2008). To improve learning from worked examples, students should be prompted to generate explanations while studying the examples (see Atkinson et al., 2000). Generating explanations aids comprehension and transfer by promoting integration of new information with prior knowledge (Chi, 2000) and by guiding attention to structural features over surface features of the to-be-learned content (McEldoon, Durkin, & Rittle-Johnson, 2013; Siegler & Chen, 2008). Asking students to generate explanations is also a recommended instructional practice for mathematics (Common Core State Standards Initiative, 2010). Thus, for each condition, we created a packet of worked examples with appropriate explanation prompts. We also included practice problems, in line with findings that worked examples should be mixed with practice problems to solve (Atkinson et al., 2000).

When designing our experimental materials, we extracted design principles from the comparison literature. First, examples to be compared were presented simultaneously to facilitate comparison (Begolli & Richland, 2015; Gentner, 1983). Second, spatial cues (e.g., side-by-side presentation) and common language were used to help students align and map the solution steps, to facilitate noticing of important similarities and differences in the examples (Namy & Gentner, 2002; Richland et al., 2007). Third, the explanation prompts focused on specific aspects of the examples to compare because this is encouraged by expert mathematics teachers (Fraivillig, Murphy, & Fuson, 1999; Huffred-Ackles, Fuson, & Sherin Gamoran, 2004; Lampert, 1990) and improves learning from comparison relative to generic, open-ended prompts "to compare" (Catrambone & Holyoak, 1989; Gentner et al., 2003).

Materials for the control condition included the same worked examples presented one at a time without supports for comparison. All students worked with a partner when studying the worked examples because working with a partner provides a familiar context for students to generate explanations, and students who collaborate with a partner tend to learn more than those who work alone (e.g., Johnson & Johnson, 1994; Webb, 1991).

Students in all conditions received the same direct instruction after studying the worked examples. We provided some direct instruction at the end of the intervention because direct instruction has been found to improve learning from comparison (Gick & Holyoak, 1983; Schwartz & Bransford, 1998; VanderStoep & Seifert, 1993).

The instructional materials often focused on multi-step equation solving. Consider the equation $3(x + 2) = 6$. Two possible first steps are to distribute the 3 (i.e., to get $3x + 6$) or to divide both sides by 3. Although the former is almost universally taught as part of the algorithm for solving this type of equation, the latter approach is arguably more efficient because it reduces the number of computations and steps needed to solve the equation. Regrettably, students often memorize rules and do not learn flexible and meaningful methods for solving equations (Kieran, 1992; Robinson & LeFevre, 2011; see also Robinson chapter in this volume). They also struggle to understand key algebraic concepts. For example, only 59% of U.S. 8th graders were able to find an equation that is equivalent to $n + 18 = 23$ (National Assessment of Educational Progress, 2011). Thus, improving students' knowledge of multi-step equation solving is greatly needed.

For student outcomes, we focused on three critical components of mathematical competence: procedural knowledge, procedural flexibility, and conceptual knowledge.

Procedural knowledge is the ability to execute action sequences to solve problems, including the ability to adapt known procedures to unfamiliar problems (Rittle-Johnson, Siegler, & Alibali, 2001). Procedural flexibility includes the knowledge of multiple methods as well as the ability to choose the most appropriate method based on specific problem features (Kilpatrick, Swafford, & Findell, 2001; Star, 2005; Verschaffel, Luwel, Torbeyns, & Van Dooren, 2009). Procedural flexibility supports efficient problem solving and is also associated with greater accuracy solving novel problems and greater understanding of domain concepts (e.g., Blöte, Van der Burg, & Klein, 2001; Carpenter, Franke, Jacobs, Fennema, & Empson, 1998; Hiebert et al., 1996). Finally, conceptual knowledge is "an integrated and functional grasp of mathematical ideas" (Kilpatrick et al., 2001). This knowledge is flexible and not tied to specific problem types, and is therefore generalizable. However, the knowledge may be implicit and not easily articulated (Alibali & Nathan, 2012; Prather & Alibali, 2009). Mathematics competence rests on students developing all three types of knowledge (Kilpatrick et al., 2001).

Studies on Comparing Methods

Our initial studies focused on comparing multiple methods for solving the same problem. Expert mathematics teachers in the U.S. (e.g. Ball, 1993; Lampert, 1990) and teachers in high performing countries such as Japan (Richland et al., 2007) often use this approach. All of the methods presented were correct, but they varied in terms of which method most appropriate and efficient for solving a particular problem. Students studied pairs of worked examples and were prompted to compare them (*compare-methods condition)* or studied the same examples one at a time and were prompted to reflect on them individually (*sequential condition*). Our comparison prompts focused student attention on recognizing that both methods adhered to domain principles, but that a particular method was more efficient for solving a particular problem.

We hypothesized that the compare-methods condition would support procedural flexibility better than the sequential condition. Comparing methods should highlight the accuracy and efficiency of multiple solution methods and facilitate knowledge and use of these methods. We also hypothesized that comparing methods would support better conceptual and procedural knowledge, as procedural flexibility is associated with both types of knowledge (e.g., Blöte et al., 2001; Carpenter et al., 1998; Hiebert et al., 1996), and reflection prompts asked students why the methods were valid.

In Rittle-Johnson and Star's (2007) study, seventh-grade students (N = 70) in pre-algebra classes learned about solving multi-step linear equations during three class periods. Students completed a packet of worked examples with their partner, explaining the procedures and answering explanation prompts; a sample worked example for each condition is illustrated in Figure 1. Before and after participating in the intervention, students completed an assessment of our three outcome measures. The procedural knowledge measure involved solving algebra equations and the conceptual knowledge measure involved recognizing or explaining algebra concepts, such as like terms indexed. Procedural flexibility was measured in two ways. The first was *use* of more efficient solution methods when solving equations; the second was knowledge of multiple ways to solve equations, including acceptance of non-standard ways to solve equations. As predicted, those who compared methods gained greater procedural flexibility. They also acquired greater procedural knowledge. The two groups did not differ in conceptual knowledge.

Students' explanations during the intervention confirmed that those who compared methods often compared the similarities and differences in solution steps across examples and evaluated their efficiency and accuracy. In turn, frequency of making explicit comparisons during the intervention predicted procedural knowledge gain.

We found similar results for 157 fifth- and sixth-grade students learning about estimating answers to multiplication problems (e.g., About how much is 37 x 29?) (Star & Rittle-Johnson, 2009). Comparing methods supported greater procedural flexibility. For students with above-average knowledge of estimation at pretest, students in the comparison condition were also more likely to maintain their gains in conceptual knowledge over a delay than students in the sequential condition. Thus, comparing methods consistently supported procedural flexibility in two domains that differed in numerous ways, including whether there was a single, correct answer and what features of the methods needed to be considered (e.g., efficiency vs. proximity to the correct answer). Its impact on procedural and conceptual knowledge was less consistent.

Theories of analogical learning help to explain how comparing methods aids learning (Gentner, 1983; Hummel & Holyoak, 1997). In both of our studies, most students were familiar with one of the solution methods at pretest. When students are familiar with one method, they can learn new methods via analogy to the familiar one. Students can make inferences about the new method by identifying its similarities and differences to a known method and making inferences about how the new method works based on its alignment with the known method. For example, students who compared methods identified how the unfamiliar methods were similar to and different from the method that they already knew; in turn these types of comparative explanations predicted learning (Rittle-Johnson & Star, 2007).

Learning via analogy to a known method requires that students are familiar with one of the methods. When one method is not familiar, an alternative learning mechanism is *mutual alignment*. During mutual alignment, people notice potentially relevant features in two unfamiliar examples by identifying their similarities and then focusing attention on and making sense of these similarities (Gentner et al., 2003; Kurtz, Miao, & Gentner, 2001). Thus, the analogy literature pointed to the potential importance of prior knowledge of one of the solution procedures when learning from comparison.

Our subsequent research explored the influence of prior knowledge on learning from comparing methods. In one study, we worked with 236 seventh- and eighth-grade students whose schools did not use a pre-algebra curriculum, and thus had had limited experience solving equations (Rittle-Johnson, Star, & Durkin, 2009). Students who did not attempt algebraic methods at pretest (i.e., novices) benefited most from studying examples sequentially, rather than from comparing methods. These students had higher procedural knowledge, conceptual knowledge and flexible use of procedures in the sequential condition than in the compare methods condition. The novices in the compare-methods condition seemed overwhelmed during the intervention – they completed less of the intervention materials and were less successful implementing non-standard methods when prompted (Rittle-Johnson et al., 2009). In contrast, students who attempted algebraic methods at pretest learned more from comparing methods; they had higher procedural knowledge and flexible use of procedures in the compare methods condition than in the sequential condition. Further, general mathematics achievement did not influence the effectiveness of comparison in this study; rather, it was prior knowledge of one of the demonstrated methods that influenced the effectiveness of comparison. These findings are consistent with cognitive load theory (Sweller, 1988); simultaneous comparison of examples has very high working memory demands, so when both methods are unfamiliar, the working memory demands may be too high and impede learning. Analogical learning can be more effective when comparing a new example to a known example, so instructors must be attentive to students' prior knowledge.

A follow-up study suggested that slowing the pace of instruction allowed novices to learn from comparing methods (Rittle-Johnson, Star, & Durkin, 2012). We worked with 198 eighth-grade students who had little prior instruction on equation solving, so we modified the materials from Rittle-Johnson et al. (2009) to cover less content in more time by focusing on fewer problem types, cutting the number of examples and explanation prompts, and adding 30 minutes to the intervention time. The impact of the compare-methods condition did not interact with use of algebra at pretest in this study. Regardless of students' prior knowledge, comparing methods supported more flexible use of procedures than sequential study, including on a one-month retention test. On other outcome measures (i.e., conceptual and procedural knowledge), the compare-methods and sequential conditions resulted in comparable learning (Rittle-Johnson et al., 2012).

In this same study, we also explored the effectiveness of delaying comparison of methods. Students studied one method on the first day, and on the second day, they compared it to alternative methods. The goal was to develop knowledge (although not mastery) of one solution method before comparing it to alternatives. However, students learned the least in this condition, relative to always comparing methods or always studying the examples sequentially. We expected delayed comparison of methods to be effective for novices and suspect that alternative instantiations of this approach could be beneficial. For example, it may be beneficial to delay comparison of multiple methods, but not to delay introducing multiple methods (e.g., initially study multiple methods sequentially, and then compare them).

There were some advantages to immediately comparing methods and no disadvantages in this study, suggesting that novices were able to learn from comparing two unfamiliar methods (i.e., mutual alignment) when the pace of instruction was slowed. Indeed, novices who compared methods often made comparisons between the two examples in their explanations, focusing on comparing problem features, solution steps, answers and the relative efficiency of the methods. Given adequate support, novices seemed able to learn by making analogies between two unfamiliar methods. However, learning via mutual alignment appears to be more difficult, and thus requires more instructional support, than learning an unfamiliar method via analogy to a known method.

Overall, comparing methods supports procedural flexibility, and sometimes conceptual and procedural knowledge as well (see Table 1 for a summary of findings from our studies). It can be used early in the learning process, but it must be carefully scaffolded (e.g., focusing on only one or two problem types and allowing ample time for study and explanation) and may result in less benefit than for students with some prior knowledge of a solution method.

Studies on Comparing Problems

Comparing correct methods is only one type of comparison. Although it is the type of comparison most often advocated for in mathematics education reforms, research on comparison in the cognitive science literature focuses on a different type of comparison – comparing different problems solved with the same method. For example, illustrating the same solution method in two stories with different cover stories, and prompting for comparison, greatly increased adults' spontaneous transfer of the solution to a new problem (Catrambone & Holyoak, 1989; Gick & Holyoak, 1983). Comparing problems is thought to support transfer by helping people abstract the key features of the method so that it is not tied to overly-narrow problem features such as the cover story. For example, adults who compared worked examples of different problems solved the same way were more likely to describe the solution to the example problems in general terms, rather than being tied to the specifics of the problem context (Catrambone & Holyoak, 1989; Gentner et al., 2003; Gick & Holyoak, 1983). In addition, explicitly stating the general method *after* learners had compared two example problems improved transfer; stating the general method without use of comparison did little to improve transfer (Gick & Holyoak, 1983). However, this previous research has been conducted with adults solving non-mathematical problems.

To explore the potential of comparing problems for helping children learn math content, we implemented two versions of comparing problems solved with the same method. In our *compare-equivalent-problems* condition, students compared two equivalent equations that varied only in the particular numbers and variables (e.g., $3(x + 2) = 6$ and $5(x + 3) = 15$); the two were solved using the same method. Prompts focused on the similarities in the solution steps and when a particular solution step could be used. In the context of solving equations, a second variation of comparing problems emerged that may better focus attention on when particular solution steps can be used. In the *compare-problem-types* condition, students compared problems with different features solved using similar, although not identical, methods. For example, they compared solutions to $3(x + 2) = 6$ and $3(x + 2) + 5(x + 2) = 16$. Prompts focused on similarities as well as differences in the solution methods due to different problem features.

In Rittle-Johnson & Star (2009), we worked with 162 seventh- and eighth-grade students who had previous experience solving equation and were randomly assigned to the compare-equivalent-problems, compare-problem-types, or compare-methods condition. Students worked on the materials with a partner during three math classes. Frequency of exposure to different solution methods was the same across conditions. We could not predict which condition would be more effective based on prior research. We found that students who compared methods gained greater procedural flexibility and conceptual knowledge than students in either compare-problem condition. Students in all three conditions made similar gains in procedural knowledge, including success transferring the methods to new problem types. These findings suggest that for mathematics learning, comparing methods supports transfer as well as comparing problems, and it supports procedural flexibility and conceptual knowledge better than comparing problems.

We also included a compare-problem-types condition in the Rittle-Johnson, Star & Durkin (2009) study described in the previous section, with students randomly assigned to a compare-problem-types, compare-methods or sequential condition. Students in the compare-problem-types condition did not differ from students in the sequential condition on any outcome, regardless of prior knowledge (Rittle-Johnson et al., 2009). Differences between the compare-problem-types and compare-methods condition depended on prior knowledge and mirrored the results for the compare-methods vs. sequential conditions. For students who attempted to use algebra at pretest, the compare-problem-types condition tended to be less effective than the compare-methods condition for procedural knowledge and flexible use of procedures. However, for students who did not use algebra at pretest, the compare-problem-types condition was more effective than the compare-methods condition for procedural knowledge, conceptual knowledge and flexible use of procedures.

Overall, our comparing problems conditions have not been especially effective in supporting learning relative to comparing methods or sequential study of examples (see Table 1). These findings corroborate concerns raised by Reed and colleagues that for complex, multi-step methods, comparing problems may not effectively support transfer (Reed, 1989; Reed, Stebick, Comey, & Carroll, 2012). Across five experiments with college students, comparing two algebra word problems and their solutions did not support learning or transfer of the solution methods. Comparing two easier arithmetic word problems supported learning of the solution method, but not transfer of the method to algebra word problems (Reed et al., 2012). In part, the lack of benefit for comparison arose because students often generated similarities and differences that were too generic or superficial to help them abstract the key features of the solution method. Complex multi-step methods may be more difficult to learn via problem comparison than simpler methods like the one learned in Gick and Holyoak (1983) for the Dunker radiation problem. Our findings also suggest that comparing problems solved with complex methods does not support procedural flexibility or conceptual understanding as well as comparing methods.

Other variations of comparing problems show more promise. First, comparing easily confusable problem types helps learners distinguish between the two problem types and solve more problems correctly (Cummins, 1992; Day, Goldstone, & Hill, 2010; VanderStoep & Seifert, 1993). For example, comparison of algebraic addition and multiplication examples supported better problem-solving accuracy than sequential study of addition examples, followed by multiplication examples (Ziegler & Stern, 2014, 2016). Second, comparing positive and negative examples of key ideas may improve conceptual knowledge. Students who compared problems that were positive and negative examples of each key idea (e.g., a line segment that was versus was not the altitude of a triangle) gained greater conceptual knowledge than students who studied only positive examples (Guo & Pang, 2011). Note that the control condition was not exposed to negative examples, making it impossible to know whether comparison was critical. Overall, comparing problems may be particularly useful in helping people recognize important problem features that differ between carefully selected problems.

Summary of Researcher-Led Classroom Studies and Proposed Guidelines

Designing studies that could be conducted in classrooms pushed our thinking about different types of comparisons, what learning outcomes each supports, and when learners are prepared to learn from comparison. We have evaluated two types of comparison in mathematics classrooms – comparing methods and comparing problems. Comparing correct methods consistently supported procedural flexibility across studies (see Table 1). For students who knew one of the solution methods at pretest, comparing methods sometimes supported greater procedural knowledge (Rittle-Johnson & Star, 2007; Rittle-Johnson et al., 2009) or greater conceptual knowledge (Rittle-Johnson & Star, 2009; Rittle-Johnson et al., 2009; Star & Rittle-Johnson, 2009). For novices, who did not know one of the solution methods at pretest, comparing methods was only helpful after we slowed the pace of the lesson. Overall, comparing methods can help a variety of students learn, but its advantages are more substantial if students have sufficient prior knowledge. How best to develop this prior knowledge is an important topic for future research.

Comparing problems solved with the same method has shown less promise for supporting mathematics learning, at least for solving equations and algebraic word problems (Reed, 1989; Reed et al., 2012; Rittle-Johnson & Star, 2009; Rittle-Johnson et al., 2009). It has been generally less effective than comparing methods, especially for students with prior knowledge in the domain. A different version of comparing problems has shown promise for helping novices learn important problem features and concepts that can be hard to learn, such as the altitude of a triangle (Guo & Pang, 2011); and future research is needed to investigate different versions of comparing problems.

Based on this research, we have begun to develop guidelines for effectively using comparison to support mathematics learning. First, *what* is compared matters. Comparing solution methods is particularly promising and can support greater math learning than other forms of comparison (Rittle-Johnson & Star, 2009). In addition, comparing a familiar method to an unfamiliar one is typically more effective than comparing two unfamiliar methods (Rittle-Johnson et al., 2009). Second, *who* is learning from comparison matters. Students with some prior knowledge in the domain often learn more effectively from comparison than novices in the domain (Rittle-Johnson et al., 2009). At the same time, once students develop strong knowledge in the domain, explicit comparison may not be needed, as students may learn equally well without comparison (Guo & Pang, 2011). This series of studies contributed significantly to a recent Practice Guide from the U.S. Department of Education (Woodward et al., 2012) that identified comparing multiple solution methods as one of five recommendations for improving mathematical problem solving in the middle grades. Thus, this series of studies has provided important information about *what* to compare and *when* in instruction to best use different types of comparison.

Year-Long Study Helping Teachers Use Comparison in Algebra I Classrooms

Given the promise of comparison for supporting mathematics learning, we developed and evaluated a set of supplementary materials for incorporating comparison throughout Algebra I instruction. Mastery of algebra is an important milestone for students, in that algebra serves as a "gatekeeper" for full participation in society and also provides students with the ability to harness new technologies and take advantage of the job opportunities resulting from them (Moses & Cobb, 2001; National Mathematics Advisory Panel, 2008). Regrettably, too few students are graduating from high school with the algebra skills needed for college or the workforce (American Diploma Project, 2004). Thus, improving algebra instruction is of great importance.

Unfortunately, U.S. teachers often struggle to effectively support comparison of multiple methods during mathematics instruction (Stein, Engle, Smith, & Hughes, 2008). Analyses of video records of mathematics instruction indicate that comparison is often not well enacted in U.S. classrooms (Richland, Holyoak, & Stigler, 2004; Richland et al., 2007). Algebra I teachers reported that they introduced multiple solution methods for at least some problem types, but that they often did not explicitly compare the methods (Lynch & Star, 2014b). These findings indicate that comparison is a reasonable adaptation of current teaching practices, but that there is a need for materials and training to help teachers effectively use comparison in mathematics instruction.

Supplemental Curriculum Materials

Based on the promise of our worked example based approach to supporting comparison, we developed a set of supplementary worked example pairs that could be used in conjunction with any Algebra I curriculum. 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 comparison materials to attempt to address them.

This led to a set of 141 worked example pairs (WEPs, available from: scholar.harvard.edu/contrastingcases). A sample WEP is shown in Figure 2. Each WEP showed the mathematical work and dialogue of two hypothetical students, Alex and Morgan, as they attempted to solve one or more algebra problems. The curriculum contained four types of WEPs, with the types varying in what was being compared and the instructional goal of the comparison, as outlined in Table 2. *Which is better?* WEPs showed the same problem solved using two correct, but different methods, with the goal of understanding when and why one method was more efficient or easier than another method for a given problem, as we did in our short-term research on comparing methods (e.g., Rittle-Johnson & Star, 2007). *Which is correct?* WEPs showed the same problem solved with a correct and incorrect method, with the goal of understanding and avoiding common errors. Comparing correct and incorrect methods supports gains in procedural knowledge, retention of conceptual knowledge, and a reduction in misconceptions (Durkin & Rittle-Johnson, 2012).

Two new comparison types were also included. The new types 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 solution methods with the goal of illuminating the conceptual rationale in one method that is less apparent in the other method. This is in contrast to the *Which is better?* comparisons, where the goal was to learn when and why one method was better for solving particular types of problems. *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. This emergence of new comparison types is one clear benefit of bridging between cognitive science and education, and highlights that the benefits are not unidirectional from theory to practice.

The WEPs were designed to maximize their potential impact on student learning based on previous research. As before, the two worked examples were presented side-by-side. To facilitate processing of the examples, we included thought bubbles where two students (Alex and Morgan) described their solution methods. We used common language in these descriptions as much as possible to help facilitate alignment of the examples. We also formalized an instructional routine to help improve the effectiveness of using our comparison materials. Each WEP had three types of reflection prompts (understand, compare, and make connections) meant to culminate in a discussion of the learning goal for the pairs of WEPs. First, *Understand* prompts, such as, "How did Morgan solve the equation?" were intended to provide students the opportunity to understand each worked example individually, prior to comparing them. Second, *Compare* prompts, such as "What are some similarities and differences between Alex's and Morgan's ways?" were meant to encourage comparison of the two worked examples. *Understand* and *Compare* prompts were very similar across WEP types and were intended to prepare students to engage in productive reflection on the final, *Make Connections* prompts, such as "On a timed test, would you rather use Alex's way or Morgan's way? Why?" and "Even though Alex and Morgan did different first steps, why did they both get the same answer?" Our pilot work revealed that sometimes teachers skipped or inadequately addressed the *Make Connection* prompts, so we supplemented each WEP with an additional "take-away" page. On this page, the fictitious students Alex and Morgan identify the learning goal for that WEP. Our intent was that the teacher would use the take-away page to provide an explicit summary statement of the instructional goal of the WEP. Prior research suggests that direct instruction is needed to supplement student-generated comparisons (Schwartz & Bransford, 1998), and a feature of high-quality instruction is that teachers summarize the instructional goals of a lesson (Brophy, 1999).

Teachers were asked to use our materials once or twice a week as a supplement to their usual curriculum. Teachers were given complete latitude in deciding which WEPs to use, when to use them, and for how long to use them.

In collaboration with Kristie Newton, we also designed a one-week, 35-hour professional development institute to familiarize teachers with the materials and approach (Newton & Star, 2013). Teachers read through and discussed the supplemental curriculum materials and viewed videotaped exemplars of other teachers using the curriculum. In addition, teachers worked in groups to plan and teach sample lessons to their peers using the materials, which were implemented and then debriefed by the group.

Implementation and Evaluation

We first piloted our materials with 13 Algebra I teachers. We then conducted a year-long randomized controlled trial that explored the effectiveness of implementing our Algebra I supplemental curriculum in typical classrooms, i.e., its impact on teachers' instruction and students' mathematical knowledge (Star, Pollack, et al., 2015). Initially, 141 Algebra I teachers from public schools were randomly assigned to either implement the comparison curriculum as a supplement to their regular curriculum (*treatment* condition) or to continue using their existing curriculum and methods ('*business as usual' control* condition). However, there was large attrition between the spring when teachers volunteered to participate and the fall when implementation began, due to a range of factors, such as teachers no longer being assigned to teach Algebra 1. Such attrition reflects the difficulties of conducting research in public high schools, but is a limitation of the study. The final sample consisted of 76 teachers and their students (39 treatment teachers with 781 students and 29 control teachers with 586 students).

Before the school year began, treatment teachers completed the 35-hour professional development institute. Observations, surveys, and interviews indicated that the professional development was successful in familiarizing teachers with our approach (Lynch & Star, 2014a). During the school year, treatment teachers implemented the materials on their own, without researcher involvement. Each time teachers used our materials, they were asked to call and report on what WEP they had used, how they had used it, and their impressions on the strengths and weaknesses of the materials. Teachers were also asked to videotape their use of our materials every other week, and teachers in both conditions were asked to videotape their instruction (without use of our materials) once a month.

Implementation results. Analysis of the videotapes indicated that treatment teachers were able to implement our materials as intended (Star, Pollack, et al., 2015). Further, control teachers did not frequently use specific instructional practices that were integral to the intervention; they rarely presented multiple methods side-by-side or explicitly compared examples. However, there was large variation in how frequently treatment teachers used our curriculum, and many treatment teachers used the supplemental materials much less frequently than intended, using our materials an average of 19 class periods during the 180-day school year (range: 0 to 56 days). In fact, 18% of participating treatment teachers did not report using the materials even once; 30% of them reported using them 5 times or fewer. Overall, treatment teachers used our supplemental materials much less often than intended (i.e., low degree of implementation), but when they chose to use the materials, it was with high fidelity (i.e., high quality of implementation).

Student outcomes. To assess student learning, teachers administered a standardized algebra readiness test, the Acuity™ Algebra Diagnostic Readiness Exam (CBT/McGraw Hill, 2007), as well as a researcher-designed assessment of algebra knowledge to their students at the beginning and end of the school year. Students in the two conditions did not differ in algebra knowledge at the beginning of the year. At the end of the school year, students' algebra knowledge was not higher in classrooms in which our materials were available (see Star, Pollack, et al., 2015 for full results). This was not surprising given how infrequently our supplemental curriculum was used in many classrooms.

Given the large variability in use of our curriculum, we evaluated whether increased use of our materials was associated with increased algebra knowledge in treatment classrooms. We calculated the *"dosage"* of curriculum given by each treatment teacher by multiplying the number of reported days each teacher used our materials by how long on average that teacher spent using our materials in a single lesson (based on the videos). Dosage ranged from 0 to 864 minutes (M = 140). We used an instrumental variable estimation (IVE) approach to analyze the data because it can account for a variety of potential factors that might teacher-level decisions about dosage (Murnane & Willett, 2011). Simply including dosage as an additional predictor in a multi-level model would create bias in the estimate of the dose-response relationship, but an IVE approach adjusts for this bias.

Increased dosage was predictive of greater procedural knowledge of algebra at the end of the school year (Star, Pollack, et al., 2015). This included solving equations, graphing equations, and factoring expressions. Dosage did not have a significant effect on conceptual knowledge, procedural flexibility or scores on the standardized algebra test. This may be because all comparison materials focused attention on procedural knowledge, whereas attention to flexibility and conceptual knowledge varied by comparison type.

We also explored whether frequency of using particular comparison types was associated with increased algebra knowledge. Frequency of using the *Why does it work?* comparisons was positively correlated with the amount of gain in students' algebra knowledge; correlations for the other types of comparison were positive but not significant.

Discussion

Overall, providing our supplemental materials to teachers did not significantly affect students' algebra knowledge. However, many teachers used our materials infrequently. Increased use of our materials had a positive effect on procedural knowledge, although not on other student knowledge outcomes. These findings suggest a potential role for comparison in supporting aspects of algebra learning, but also point to the challenges of supporting teachers' integration of this approach into the curriculum.

The largest limitation was that teachers used our materials much less than intended. Almost half of the sample reported using our materials on 5 or fewer occasions across the entire school year; on average, teachers used our materials for 19 class periods for a total of 140 minutes across the entire school year. We intentionally chose to give teachers a great deal of choice in which materials to use and when to use them, but this freedom seemed to lead to a substantial number of teacher not using the materials or using them only infrequently. Interest, training, and carefully designed materials were not enough for many teachers to adopt our materials with much regularity. It is time consuming and potentially challenging to select, plan, and integrate supplemental materials with the existing curriculum. Teachers would likely benefit from more clear structure and direction for which materials to use when.

Further, teachers could use additional supports to implement higher-quality instruction. Too frequently in our study, teachers were providing the explanations, with little student explanation or discussion (Star, Newton, et al., 2015; Star, Pollack, et al., 2015). In addition, our professional development only occurred over the summer, but effective mathematics teacher professional development typically provides teachers with multiple, intensive opportunities to establish meaningful connections between existing instructional practices and the practices and beliefs advocated by the new approach across the school year (Garet, Porter, Desimone, Birman, & Yoon, 2001). More sustained professional development is likely needed to increase high-quality instruction that emphasizes student explanation and discussion when integrating our materials.

Conclusions

Comparison is a promising instructional approach for supporting mathematics learning. In short-term classroom studies, comparing different solution methods for solving the same problem often promoted procedural flexibility and sometimes promoted conceptual and procedural knowledge. Further, because comparison requires substantial cognitive effort, it is often more effective when learners have some prior domain knowledge; novices can become overwhelmed by comparison without adequate support. In addition, comparison can be infused throughout curriculum materials for an entire course, and some teachers are able to implement comparison-based materials in their instruction. However, providing comparison-based materials and professional development in the summer was not sufficient for many teachers to incorporate comparison throughout the school year or to reliably improve student outcomes. A variety of changes to the supplemental materials and professional development may be needed.

Future research is needed to advance both theory and practice. First, research is needed on how different types of comparison support different learning outcomes. For example, the *Why does it work?* comparison type that emerged from our classroom research is particularly promising for supporting knowledge of critical math concepts and procedures. Do *Why does it work?* comparisons support greater student learning than sequential study of the same materials? We predict that they will. More generally, theories of analogical learning need to explicitly consider what is being compared. For example, structure-mapping theory (Gentner, 1983, 2010) should be refined to consider how comparing problems vs. methods impacts alignment and mapping of elements between two examples. Second, research is needed on the supports teachers need to effectively integrate comparison in their instruction, including the features of the supplemental materials (e.g., use of worked examples vs. student-generated solutions) and the importance of student explanation and discussion. Comparison is an effective practice for improving mathematics learning, and future research should further refine theory and practice to maximize its potential benefits.

References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparison: A meta-analytic review. *Educational Psychologist, 48*(2), 87-113. doi:10.1080/00461520.2013.775712

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in Mathematics Teaching and Learning: Evidence From Learners' and Teachers' Gestures. *Journal of the Learning Sciences, 21*(2), 247-286. doi:10.1080/10508406.2011.611446

American Diploma Project (2004). Ready or Not: Creating a High School Diploma That Counts. Achieve Inc. Retrieved from http://www.achieve.org/files/ReadyorNot.pdf August 2014.

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*(2), 181-214. doi:10.2307/1170661

Australian Education Ministers (2006). *Statements of learning for mathematics*. Curriculum Corporations. Retrieved from http://www.curriculum.edu.au/verve/_resources/SOL_Mathematics_2006.pdf August 2014.

Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *The Elementary School Journal, 93*, 373-397. doi:10.1086/461730

Begolli, K. N., & Richland, L. E. (2015). Teaching Mathematics by Comparison: Analog Visibility as a Double-Edged Sword. *Journal of Educational Psychology*. doi:10.1037/edu0000056

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. doi:10.1037//0022-0663.93.3.627

Brophy, J. (1999). Teaching *Education Practices Series No. 1, International Bureau of Education*. Geneva.

Carpenter, T. P., Franke, M. L., Jacobs, V. R., Fennema, E., & Empson, S. B. (1998). A longitudinal study of invention and understanding in children's multidigit addition and subtraction. *Journal for Research in Mathematics Education, 29*(1), 3-20. doi:10.2307/749715

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(6), 1147-1156. doi:10.1037//0278-7393.15.6.1147

CBT/McGraw Hill. (2007). Acuity Algebra. Retrieved from http://www.acuityforschools.com

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* (Vol. 5, pp. 161-238).

Cummins, D. (1992). Role of analogical reasoning in the induction of problem categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 1103-1124. doi:10.1037//0278-7393.18.5.1103

Day, S., Goldstone, R. L., & Hill, T. (2010). *The effects of similarity and individual differences on comparison and transfer.* Paper presented at the Cognition in Flux: Proceedings of the 32nd Annual Meeting of the Cognitive Science Society, Austin, TX.

Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction, 22*(3), 206-214. doi:10.1016/j.learninstruc.2011.11.001

Fraivillig, J. L., Murphy, L. A., & Fuson, K. (1999). Advancing Children's Mathematical Thinking in Everyday Mathematics Classrooms. *Journal for Research in Mathematics Education, 30*, 148-170.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915-945. doi:10.3102/00028312038004915

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science: A Multidisciplinary Journal, 7*(2), 155-170. doi:10.1016/S0364-0213(83)80009-3

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*(5), 752-775. doi:10.1111/j.1551-6709.2010.01114.x

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*(2), 393-405. doi:10.1037/0022-0663.95.2.393

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), 1-38. doi:10.1016/0010-0285(83)90002-6

Goldstone, R. L., Day, S., & Son, J. (2010). Comparison. In B. Glatzeder, V. Goel, & A. von Müller (Eds.), *On thinking: Towards a theory of thinking* (Vol. II, pp. 103-122). Heidelberg, Germany: Springer Verlag GmbH.

Guo, J.-p., & 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. doi:10.1007/s10212-011-0060-y

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K. C., Human, P., Murray, H., . . . Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher, 25*(4), 12-21. doi:10.2307/1176776

Huffred-Ackles, K., Fuson, K., & Sherin Gamoran, M. (2004). Describing levels and components of a math-talk learning community. *Journal for Research in Mathematics Education, 35*(2), 81-116. doi:10.2307/30034933

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*(3), 427-466. doi:10.1037/0033-295X.104.3.427

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

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). New York: Simon & Schuster.

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

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

Kurtz, K., Miao, C.-H., & Gentner, D. (2001). Learning by analogical bootstrapping. *The Journal of the Learning Sciences, 10*, 417-446. doi:10.1207/S15327809JLS1004new_2

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. doi:10.2307/1163068

Lynch, K., & Star, J. R. (2014a). Exploring teachers' implementation of comparison in Algebra I. *Journal of Mathematical Behavior, 35*, 144-163. doi:10.1016/j.jmathb.2014.07.003

Lynch, K., & Star, J. R. (2014b). Teachers' Views about Multiple Strategies in Middle and High School Mathematics. *Mathematical Thinking and Learning: An International Journal, 16*(2), 85-108.

McEldoon, K. L., Durkin, K. L., & Rittle-Johnson, B. (2013). Is self-explanation worth the time? A comparison to additional practice. *British Journal of Educational Psychology, 83*(4), 615-632. doi:10.1111/j.2044-8279.2012.02083.x

Moses, R. P., & Cobb, C. E. (2001). *Radical Equations: Math Literacy and Civil Rights*. Boston: Beacon Press.

Murnane, R. J., & Willett, J. B. (2011). *Methods matter: Improving causal inference in educational and social science research*. New York: Oxford University Press.

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. doi:1O.1037//O096-LWS1.31.I.5

National Assessment of Educational Progress. (2011). Mathematics Performance. Retrieved from http://nces.ed.gov/programs/coe/indicator_cnc.asp

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

Newton, K. J., & Star, J. R. (2013). Exploring the nature and impact of model teaching with worked example pairs. *Mathematics Teacher Educator, 2*, 86-102. doi:10.5951/mathteaceduc.2.1.0086

Newton, K. J., Star, J. R., & Lynch, K. (2010). Understanding the Development of Flexibility in Struggling Algebra Students. *Mathematical Thinking and Learning: An International Journal, 12*, 282-305. doi:10.1080/10986065.2010.48215

NCTM. (2014). *Principles to Actions: Ensuring Mathematical Success for All*. Reston, VA: National Council of Teachers of Mathematics, Inc.

Oakes, L. M., & Ribar, R. J. (2005). A comparison of infants' categorization in paired and successive presentation familiarization tasks. *Infancy, 7*(1), 85-98. doi:10.1207/s15327078in0701_7

Prather, R. W., & Alibali, M. W. (2009). The Development of Arithmetic Principle Knowledge: How Do We Know What Learners Know? *Developmental Review, 29*(4), 221-248.

Reed, S. K. (1989). Constraints on the abstraction of solutions. *Journal of Educational Psychology, 81*(4), 532-540. doi:10.1037/0022-0663.81.4.532

Reed, S. K., Stebick, S., Comey, B., & Carroll, D. (2012). Finding similarities and differences in the solutions of word problems. *Journal of Educational Psychology, 104*(3), 636-646. doi:10.1037/a0027181

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

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

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*(2), 346-362. doi: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. doi:10.1037/0022-0663.99.3.561

Rittle-Johnson, B., & Star, J. R. (2009). Compared to what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology, 101*(3), 529-544. doi: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. doi: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*, 436-455. doi:10.1111/j.2044-8279.2011.02037.x

Robinson, K. M., & LeFevre, J.-A. (2011). The inverse relation between multiplication and division: Concepts, procedures, and a cognitive framework. *Educational Studies in Mathematics, 79*(3), 409-428. doi:10.1007/s10649-011-9330-5

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*(4), 475-522. doi:10.1207/s1532690xci1604_4

Siegler, R. S., & Chen, Z. (2008). Differentiation and Integration: Guiding Principles for Analyzing Cognitive Change. *Developmental Science, 11*(4), 433-448.

Singapore Ministry of Education. (2006). Secondary Mathematics Syllabuses.

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36*, 404-411.

Star, J. R., Newton, K., Pollack, C., Kokka, K., Rittle-Johnson, B., & Durkin, K. (2015). Student, teacher, and instructional characteristics related to students' gains in flexibility. *Contemporary Educational Psychology, 41*, 198-208. doi:10.1016/j.cedpsych.2015.03.001

Star, J. R., Pollack, C., Durkin, K., Rittle-Johnson, B., Lynch, K., Newton, K., & Gogolen, C. (2015). Learning from comparison in algebra. *Contemporary Educational Psychology, 40*, 41-54. doi:10.1016/j.cedpsych.2014.05.005

Star, J. R., & Rittle-Johnson, B. (2008). Flexibility in problem solving: The case of equation solving. *Learning and Instruction, 18*, 565 - 579. doi:10.1016/j.learninstruc.2007.09.018

Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology, 101*, 408 - 426. doi:10.1016/j.jecp.2008.11.004

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. doi:10.1080/10986060802229675

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257-285. doi:10.1016/0364-0213(88)90023-7

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*(1), 59-89. doi:10.1207/s1532690xci0201_3

Treffers, A. (1991). Realistic mathematics education in the Netherlands 1980-1990. In L. Streefland (Ed.), *Realistic Mathematics Education in Primary School*. Utrecht, the Netherlands: Freudenthal Institute, Utrecht University.

VanderStoep, S. W., & Seifert, C. M. (1993). Learning "how" versus learning "when": Improving transfer of problem-solving principles. *Journal of the Learning Sciences, 3*(1), 93-111.

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), 335-359. doi: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.

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

Ziegler, E., & Stern, E. (2014). Delayed benefits of learning elementary algebraic transformations through contrasted comparisons. *Learning and Instruction, 33*, 131-146. doi:10.1016/j.learninstruc.2014.04.006

Ziegler, E., & Stern, E. (2016). Consistent advantages of contrasted comparisons: Algebra learning under direct instruction. *Learning and Instruction, 41*, 41-51. doi:10.1016/j.learninstruc.2015.09.006

Table 1

Design Features and Outcomes From Our Experimental Studies on Using Comparison to Support Mathematics Learning

	Rittle-Johnson & Star (2007)	Star & Rittle-Johnson (2009)	Rittle-Johnson, Star & Durkin (2009)	Rittle-Johnson, Star & Durkin (2012)	Rittle-Johnson & Star (2009)
Design Features					
Instructional Conditions	Compare-methods or sequential	Compare-methods or sequential	Compare-methods, sequential or compare-problem-types	Compare-methods, sequential, or delayed-compare-methods	Compare-methods, compare-equivalent-problems or compare-problem-types
Target Task	Linear equations	Computational Estimation	Linear equations	Linear Equations	Linear equations
Most children familiar with a target method at pretest	Yes	Yes	Mixed	Mixed	Yes
Condition(s) with highest performance					
Conceptual Knowledge	Same (poor measure)	Depends on prior knowledge for retention	Depends on prior knowledge	Same	Compare-methods
Procedural Knowledge	Compare-methods	Same	Depends on prior knowledge	Similar	Same
Flexibility	Compare-methods	Compare-methods	Depends on prior knowledge	Compare-methods	Compare-methods

Table 2

Four Types of Comparisons Used in Our Year-Long Algebra I Classroom Study

	Comparison Type			
Comparison goal	*Which is better?*	*Which is correct?*	*Why does it work?*	*How do they differ?*
Features of examples	Same problem solved with two different, correct methods.	Same problem solved with a correct and an incorrect method.	Same problem solved with two different, correct methods.	Different problems. Focus is not on solution methods.
Focus of comparison	When and why a method is more efficient or easier.	Why one method works and one does not.	Conceptual rationale revealed in one method that is less apparent in other method.	What relations between problems and answers reveal about underlying concept.
Sample worked example pair	Solving the proportion $4/5 = 24/n$ by finding equivalent fractions or by cross-multiplying.	Solving $45y + 90 = 60y$ by first subtracting $45y$ from both sides or by incorrectly combining $45y + 60$.	Expanding the expression (x^4)^2 by applying the power rule (x ^(4*2)) or by expanding and then squaring.	Graphing the equations $y = x^2$ and $y = -x^2$.
Sample *Make Connections* prompt	Can you make up a general rule for when Alex's way is better and when Morgan's way is better?	Can you state a general rule about combining like terms?	Even though Alex and Morgan did different first steps, why did they get the same answer?	How does changing the sign of the coefficient of x^2 affect the graph of the quadratic function?

A. Compare-Methods Condition

Jessica's Solution:		**Mary's Solution:**	
$2(t - 1) + 3(t - 1) = 10$		$2(t - 1) + 3(t - 1) = 10$	
$2t - 2 + 3t - 3 = 10$	______________	$5(t - 1) = 10$	______________
$5t - 5 = 10$	*Combine*	$t - 1 = 2$	*Divide on Both*
$5t = 15$	*Add on Both*	$t = 3$	*Add on Both*
$t = 3$	*Divide on Both*		

Label the first step for each solution in the blank space provided above.

1. Jessica and Mary did different first steps. Is it OK to do either step first in this problem? Explain your reasoning.

2. Why might it be helpful to know two different ways to solve equations like this one?

B. Sequential Condition

Jessica's Solution:	
$2(t - 1) + 3(t - 1) = 10$	
$2t - 2 + 3t - 3 = 10$	______________
$5t - 5 = 10$	*Combine*
$5t = 15$	*Add on Both*
$t = 3$	*Divide on Both*

Label the first step in the blank space provided above.

1. Why did Jessica divide as her last step?

-----NEXT PAGE-----

Mary's Solution:	
$4(t - 5) + 3(t - 5) = 14$	
$7(t - 5) = 14$	______________
$t - 5 = 2$	*Divide on Both*
$t = 7$	*Add on Both*

Label the first step in the blank space provided above.

1. Do you think the solution method used on this problem is a good one? Why?

Figure 1. Sample pages from the compare-methods and sequential conditions from Rittle-Johnson & Star (2007).

Why does it work?

Alex and Morgan were asked to simplify $\frac{7}{a} \div \frac{b}{c}$

Alex's "divide first" way	Morgan's "multiply by the reciprocal" way
$\frac{7}{a} \div \frac{b}{c}$	$\frac{7}{a} \div \frac{b}{c}$
↓	↓
$\dfrac{\frac{7}{a}}{\frac{b}{c}}$	$\frac{7}{a} \cdot \frac{c}{b}$
First I rewrote the division problem in fraction notation.	**First I rewrote the problem as multiplication by the reciprocal of *b*/*c*, which is *c*/*b*.**
↓	↓
$\dfrac{\frac{7}{a} \cdot \frac{c}{b}}{\frac{b}{\not c} \cdot \frac{\not c}{b}}$	$\frac{7c}{ab}$
Then I multiplied the numerator and the denominator by the reciprocal of the denominator, *c*/*b*. The terms in the denominator canceled out.	**I multiplied the terms in the numerator and denominator together, and I got my answer.**
↓	
$\frac{7c}{ab}$	
I multiplied the terms in the numerator to get my answer.	

*** How did Alex simplify the expression?**
*** How did Morgan simplify the expression?**
*** Why do the terms in the denominator cancel out in Alex's second step?**
*** What are some similarities and differences between Alex's and Morgan's ways?**
*** Even though Alex and Morgan did different first steps, why did they both get the same answer?**

11.2.2

Figure 2. Sample worked example pair from our Algebra I scale-up project, illustrating a *Why does it work?* comparison (Star, Pollack, et al., 2015).