PROCEDURAL VERSUS CONCEPTUAL ALGEBRA REFRESHER INTERVENTIONS AND EFFECTS OF COGNITIVE ABILITIES IN ADULTS

A Master's Thesis

Presented to

The Faculty of the Department

Of Psychology

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Arts

By

John T. Elias

December, 2012

PROCEDURAL VERSUS CONCEPTUAL ALGEBRA REFRESHER INTERVENTIONS

AND EFFECTS OF COGNITIVE ABILITIES IN ADULTS

John T. Elias

APPROVED:

Paul Cirino, Ph.D.
Committee Chair

Carla Sharp, Ph.D.

Paul Massman, Ph.D.

John W. Roberts, Ph.D.
Dean, College of Liberal Arts and Social Sciences
Department of English

PROCEDURAL VERSUS CONCEPTUAL ALGEBRA REFRESHER INTERVENTIONS AND EFFECTS OF COGNITIVE ABILITIES IN ADULTS

An Abstract of a Master's Thesis

Presented to

The Faculty of the Department

Of Psychology

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Arts

By

John T. Elias

December, 2012

ABSTRACT

The success rate of high school and college students in algebra is low (Tyson et al., 2007; Croft, 2006). As such, the present study compared the effectiveness of conceptual and procedural algebra "refresher" interventions for 63 college students and investigated the relationship of cognitive variables to math outcome. Results revealed that participants in both interventions improved significantly from pretest to posttest, but that there was no significant difference between the two interventions. However, a follow-up analysis in which participants who scored at ceiling or at floor on pretest were excluded revealed an advantage for the conceptual treatment group relative to the procedural group in terms of educationally meaningful effect size ($d = +0.53$). There were no relationships between cognitive variables and math outcomes; however, in the follow-up analysis, long-term memory was positively correlated with posttest performance, but working memory and executive functioning were still unrelated to math outcomes. There was no interaction between cognitive variables and posttest performance of the two groups, which likely reflects the lack of differential treatment effects or robust zero-order correlations. The present study addresses the shortage of research on algebra, particularly the shortage of experiments that compare the effectiveness of different algebraic interventions. It also provides insight into intervention methods that educators may potentially use.

Keywords: algebra interventions, math difficulty, long-term memory, working memory, executive functions

TABLE OF CONTENTS

PROCEDURAL VERSUS CONCEPTUAL ALGEBRA REFRESHER INTERVENTIONS IN ADULTS AND MODERATING EFFECTS OF COGNITIVE ABILITIES

Introduction

Prevalence of Difficulty in Math

Learning math is difficult. In a Gallup survey of more than one thousand American children ages thirteen to seventeen, thirty seven percent of respondents reported that math was their most difficult subject in school, more than that who reported science (20%) or English (18%) to be their most difficult subject (Saad, 2003).

With regard to research in the field of academic learning, the math research literature is not as mature as the reading literature, and as of yet, is not as extensive (Chiappe, 2005), despite rapid growth in this area, particularly over the last decade. In addition, in educational settings, greater emphasis is placed on reading relative to math. For example, teachers in early education tend to focus more on language than math, and the amount of time and resources parents of young children tend to invest in providing math context is considerably less than the time and resources invested in providing language context (Cannon & Ginsburg, 2008). There is considerable consensus on underlying deficits that lead to reading difficulty—for example, deficits in phonological awareness and rapid naming of letters (Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004)—and as a consequence, a number of empirically-derived reading interventions and preventative measures have been developed. Accordingly, government-supported reading interventions such as Reading First (U.S. Department, 2009), Early Reading First (U.S. Department, 2012), No Child Left Behind (U.S. Department, 2010), and Get Ready to Read (2011) have been heavily funded.

In contrast, much work remains until we achieve a similar understanding of all the variables that play a causal role in math difficulties; as such understanding increases, preventative and intervention procedures are likely to follow (Chiappe, 2005). The Center for Education (2001) and the National Math Advisory Panel (NMAP) (2008) conceded that, albeit the great strides in the area of math learning research as of late, the field still has room to grow. While U.S. schools have seen success in implementing reading interventions, math continues to be problematic, and the United States continues to lag behind other industrialized countries in mathematics. The Center stressed the global importance of learning mathematics and of understanding why mathematics works the way it works, emphasizing that “All young Americans must learn to think mathematically, and they must think mathematically to learn.” (Center for Education, p. 16). Of particular relevance for the present study, in their recommendations, they addressed the particular need for research into instructional strategies that would assist individuals in transitioning from arithmetic to algebraic thinking.

Math Education and Performance

Math education in the United States has followed a cyclical process, in terms of the teaching emphasis and the motivation for learning (Miller & Mercer, 1997). The emphasis was on learning basic math for everyday living in the early 1900s and shifted to an accentuation of discovery methods in the 1930s. The 1950s saw an increased emphasis of discovery methods and concern over math performance at a time when mathematical and technological prowess was deemed exigent in establishing the nation’s world status after losing the space race to the Soviet Union. Finally, in recent decades, math education returned to emphasize fundamental basics when it became evident that students with math learning disabilities were being left behind (Miller & Mercer, 1997). There is, however, considerable variability in the approach currently taken toward math instruction, with some districts and curricula focused more on discovery learning, with others focused on more explicit instruction, including practice with basic facts (Star & Rittle-Johnson, 2008; NMAP, 2008).

The current state of math performance is low, especially in the U.S, and especially for higher-level math, despite small recent increases in the percentage of students at or above the "proficient" level in math according to the National Assessment of Educational Progress (NAEP, 2011). Sixty percent of fourth graders and 65% of eighth graders perform lower than a proficient level. Furthermore, 18% of fourth graders and 27% of eighth graders perform below even a basic, fundamental level. These numbers show a trend of decreasing proficiency with age, especially as required math concepts become difficult (NAEP, 2011).

Arithmetic and early foundation-level math have been the focal point for much of the math research thus far, in regards to both descriptive and intervention studies. However, beyond arithmetic, mathematics encompasses a range of skills. Of these, algebra has received attention, given its role as a bridge between basic and more advanced mathematics. Students who struggle with algebra have difficulty progressing through more advanced mathematical courses, particularly for science, technology, engineering, or mathematics (STEM) based careers (NMAP, 2008). Therefore, learning algebra early on can be beneficial for those interested in pursuing higher education. High school students who take and pass algebra and increasingly higher-level math courses are more likely to graduate from college, and of those who graduate from college, those who learn higher-level math in high school are more likely to graduate with a STEM degree (Tyson et al., 2007). College algebra is useful for preparation for STEM degrees but is increasingly becoming a requirement for non-STEM degrees as well; fewer than 20% of students taking college algebra plan to major in a STEM subject (Herriott & Dunbar, 2009). Even beyond degree or career requirements, advanced mathematical knowledge is still relevant for competitiveness in a culture that is increasingly technological (Center for Education, 2001). Learning algebraic concepts is useful for even everyday living activities such as planning purchases, renovating a house, or calculating interest on investments. Its mastery is mandatory for many undergraduate degrees, and it is essential for those wishing to pursue a career in STEM.

The content of algebra is markedly different in character from the basic math usually taught before it, and it can, hence, be especially difficult to learn. According to Tyson et al. (2007) 32.8% of individuals in high school either do not take or do not pass any math class Algebra I or higher. Difficulty in learning algebra is not unique to high school students; the success rate in college algebra tends to be much lower than in other freshman classes such as English, government, economics, and biology. Grade distributions at two 2-year institutions indicated that 47.3% and 63% of students received a D, F, or withdrew from college algebra courses (Herriott & Dunbar, 2009). According to Croft (2006), as much as 17% of the total undergraduate population at the University of Houston take introductory college algebra in a given year, and as many as 22.2% of those students fail each year, illustrating the need for further investigation into how adults learn algebra.

Cognitive Functions and Math

Much of the research relating cognitive constructs to math has focused on elementary math processes such as quantity estimation, number sense, place value, simple arithmetic, fact retrieval, and simple problem solving (e.g, Cirino, 2011; Fuchs et al., 2009; Geary, Hoard, & Hamson, 1999) with only a few studies focused on cognitive skills involved in algebra or higher-level math processes (e.g., Tolar, Lederberg, & Fletcher, 2009). Conclusions across studies vary but generally agree that verbal ability, memory, executive functioning, and spatial cognition are important skills for elementary math processes (e.g., Geary, Hoard, Nugent, & Byrd-Craven, 2007; Swanson & Kim, 2007; Geary et al., 2009). One study of third grade students showed that verbal reasoning, verbal concepts, nonverbal reasoning, planning, visuospatial working memory, phonological working memory, and verbal working memory were underlying factors in recalling simple rules about numbers, computing addition and subtraction problems, and solving word problems (Männamaa et al., 2012). In a meta-analysis, Swanson and Jerman (2006) found that average-achieving children performed better than children with MD on different measures of problem solving, working memory, and long term memory and that children with MD performed better on these tasks than children with comorbid MD and reading disability (MDRD). A similar meta-analysis by Swanson, Jerman, and Zheng (2009) saw that those with reading disabilities (RD) performed worse on measures of working memory and short term memory than average achievers. The two main methods of investigating the determinants of math outcomes include correlational studies (relating cognitive abilities to math outcomes across a continuum of performance) or causal-comparative studies (comparing pre-existing groups, such as individuals with or without math disability. However, there is a deficiency of studies that examine cognitive constructs in the context of true experiments with randomization of subjects, as these skills may moderate treatment effects in this context.

Memory

Memory is among the most studied of cognitive abilities with regard to conceptualizing mathematics learning, and within memory, working memory (WM) is the component that has received the most attention. According to Cowan (2010), WM is the ability to hold information in an accessible state in order to complete cognitive tasks. A number of studies indicate that executive and spatial aspects of WM impairment tend to be associated with mathematics disability (MD) (McLean & Hitch, 1999; Hitch & McAuley, 1991; Siegel & Ryan, 1989; Adams

& Hitch, 1997; Holmes & Adams, 2006). DeCaro, Rotar, Kendra, and Beilock (2010) proposed that interventions which reduce situation-induced worries free up the resources on which individuals with high WM tend to rely, suggesting that those with high WM may benefit the most from such interventions. On the other hand, Huang-Pollock and Karalunas (2010) showed that skill acquisition was more negatively affected for those with low WM than with high WM when demands were increased, suggesting that an intervention which reduced WM demand would be most helpful for those with low WM. Measures of WM usually involve either (1) storing information while performing a less pertinent task; (2) storage and mental manipulation of that stored information; or (3) tracking a continuous sequence according to the relationship among its elements. "Digits" tasks, which require a subject to listen to a string of numbers and repeat it backward are perhaps the most common WM exemplar, and satisfy the requirements under (2) above; such tasks are robust in their relation to math performance (e.g., Bull, Espy, & Wiebe, 2008). Digits tasks are also part of many commonly used batteries that test WM such as the Working Memory Test Battery for Children (WMTB-C) (Pickering & Gathercole, 2001) or the Weschler Adult Intelligence Scale (WAIS) (Weschler, 1997).

Beyond WM, the relationship between long-term memory (LTM) and math has also been investigated; for example, Prevatt, et al. (2010) found that LTM was related to math applications, computations, and concepts, but not to math fluency. Seethaler and Fuchs (2006) suggest that verbal LTM may be more related to word problems than computational estimation. Both of these studies use a subtest of the Woodcock Johnson-III Cognitive Battery (WJ-III) that requires memory of already-acquired information such as animal names (*Retrieval Fluency*) as a measure of LTM. The study by Prevatt, et al. (2010) used the WJ-III *Visual-Auditory Learning* subtest which requires learning symbol meanings and then reading sentences formed from these symbols. However, LTM measures differ in their demands from one another as well as from other commonly used tasks. Therefore, it would be beneficial to understand the relationship of more types of LTM to math. For example, in list-learning tasks, participants are repeatedly presented with a list until mastery. Such tasks can be used in assessment of both immediate and delayed verbal recall. List-learning tasks provide repeated exposure that requires self-organization and retention over time. The relation of LTM and math is understudied and will become more cogent as more forms of LTM are included in analyses that seek to understand its relationship to math.

In contrast to a robust literature on the cognitive determinants of arithmetic, less is known about the determinants of higher-level math processes such as algebra. However, recent studies have highlighted the relationships of WM and LTM (to a greater and lesser degree, respectively), to algebra. In a study of college students, computational fluency, spatial visualization, and WM all had an effect on algebra achievement (Tolar et al., 2009). Furthermore, computational fluency and spatial visualization were shown to be mediators of the effect of WM on algebra achievement. Another study (Lee et al., 2011) showed that strengths in pattern-making and procedural computation were predictive of algebraic performance and that “updating,” which involves maintaining information while simultaneously absorbing new information, was predictive of both pattern-making and procedural computation in algebra. In a different large-scale study of 11-year-olds, WM explained 23 and 27% of the variance in representation formation and solution formation of algebraic word problems, respectively (Lee, Ng, & Ng, 2009). That WM is important for algebra is not surprising given the relationship of WM to math in general and the demands of algebra specifically. For example, algebraic word problems require the ability to represent words with numbers and also to choose the correct operators.

Although there is growing evidence for the role of WM for even older learners of more complex math like algebra, less is known about other types of memory such as LTM. Adults with low levels of math education who had taken algebra and geometry several years previous showed a pattern of steep and steady decline in LTM of algebra and geometry over the course of many years, whereas individuals with moderate to high levels of math education showed very little forgetting (Bahrick & Hall, 1991). This relationship could be due to a number of factors, including but not limited to stronger encoding of the algebraic information, better contextualization, or deeper processing of the information.

Executive Functioning

Another important construct used to conceptualize mathematics learning is executive functioning (EF). Working memory is sometimes included as a component of EF, though is treated separately above given its prominence in the mathematical literature and the robust empirical and theoretical literature on WM specifically (e.g., Engle, 2002). EF can be defined in numerous ways, though commonly it includes the management, organization, planning, and direction of other higher-level processes (Lee et al., 2011). A number of studies highlight the relationship of EF to math outcomes. For example, Mazzocco and Kover (2007) found that inhibitory control at ages 6 and 7 was a significant predictor of mathematics performance in later elementary school years. Bull, Espy, and Wiebe (2008) found corroborating evidence, showing that inhibitory control at age 7 was predictive of math performance and found that planning and monitoring skills were predictive as well. A study of kindergarten students showed that inhibitory control and attention shifting were associated with numeracy, shape knowledge, quantity, addition, and subtraction (Blair & Razza, 2007). Further evidence of the relationship between EF and math is exemplified in findings that students with math disabilities performed significantly lower than those without math disabilities on a measure of planning and monitoring (Sikora, Haley, Edwards, & Butler, 2002). While much of the research in this area has studied children, a study by Osmon, Smerz, Braun, and Plambeck (2006) showed that EF in college adults was also related to math achievement. Of the many components of EF such as inhibition, planning, organizing, shifting attention, monitoring (Lee et al., 2011), planning is one key component of EF that may be especially important in higher-level math because of the multistep nature of algebra.

Other Predictors of Math

LTM, WM, and EF are strong cognitive predictors of math, but other variables beyond specific cognitive processes also contribute to math performance, so inclusion of these other variables is useful to help contextualize the roles of LTM, WM, and EF. These variables include initial math ability (Geary, 1993; Gersten, Clarke, & Mazzocco, 2007), reading ability (Jiban & Deno, 2007), and math anxiety and motivation (Wigfield & Meece, 1988; Wang & Pomerantz, 2009).

Initial Math Ability

Given the complexity of math, it is not surprising that individual differences are robust. Some individuals have strong math skills, but others have math skills weak enough to be classified as a math learning disability depending on the cut-point used to define them; math disabilities affect anywhere between 5.9% (Barbaresi, Katusic, Colligan, Weaver, & Jacobsen, 2005) and 17% (Mazzocco & Myers, 2003) of individuals. Math ability varies along a continuum, and one's underlying math ability is predictive of performance on tasks which test specific math outcomes (Tolar et al., 2009). This is especially likely for adults, but in the context of a randomized study, inclusion of pretest math skill should increase power for detecting group differences (Shadish, Cook, & Campbell, 2002).

Reading

Reading ability is closely tied to areas of math ability. For example, Jiban and Deno (2007) found that for statewide tests in Minnesota, performance in reading was significantly predictive of math performance at both third (R^2 = 17%) and fifth (R^2 = 38%) grades. Another study involving third graders showed that language skills and reading skills were significantly related to addition and subtraction estimation skills (Seethaler & Fuchs, 2006). A number of studies involving children with concurrent math and reading disabilities (MDRD) show that children with MDRD perform significantly worse than children with only a math disability on story problems in particular (e.g., Jordan & Hanich, 2000; Powell, Fuchs, Fuchs, Cirino, & Fletcher, 2008; Jordan & Montani, 1997), and generally across a range of other math skills as well (Cirino, Fuchs, Elias, Powell, & Schumacher, submitted). Because of the linguistic and textual nature of story problems, it makes sense that reading ability might be tied to math story problem ability (Powell et al., 2008). Lager (2006) showed that reading and linguistic ability were related to performance on algebra-related tasks for tasks that were more language-intensive or less language-intensive. Due to its close ties with math ability, reading ability is a useful construct to include in analyses in which math is an outcome.

Math Perception: Anxiety and Motivation

A number of studies indicate that math anxiety is common among children (Wigfield & Meece, 1988; Gierl & Bisanz, 1995) and adults (Prevatt et al., 2010; Hendel, 1980) and has recently been shown to have a distinct neural signature (Young, Wu, & Menon, 2012). The reasons math provokes greater anxiety than other academic subjects are unclear, but they could be related to a number of different factors. For example, lack of support from instructors with high expectations has been demonstrated to elicit anxiety (Turner et al., 2002). Furthermore, the importance an individual places on math achievement is inversely related to the level of math anxiety, i.e., greater perceived importance is associated with less anxiety (Meece, Wigfield, & Eccles, 1990). The perception of a higher working memory (WM) demand for math than for other academic subjects increases actual difficulty and subsequently provokes anxiety (Hopko et al., 1998; DeCaro et al., 2010).

Motivation is also important to consider with regard to math learning. Results of a longitudinal study of 825 American and Chinese middle-school-age children indicated that the quality of children’s motivation for mastery decreases with age, and this decrease is especially marked in the early teen years. The value placed on academics as a whole was a motivator and protective factor for children in China, whereas the lower value American children placed on academics accompanied declines in motivational behavior (Wang & Pomerantz, 2009). According to Zimmerman, Bandura, and Martinez-Pons (1992), academic motivation stems from personal goal setting and perceived academic efficacy. Students that feel they are capable of self-regulating their activities are more confident of their ability to master academic concepts, are more motivated to attain academic goals, and ultimately perform better academically.

IQ as a Predictor of Math

The measurement of intelligence varies, but it is clear that batteries of tests which produce an “IQ” score do correlate with a wide variety of cognitive measures, likely because they themselves are often included in the battery. Francis, Fletcher, Shaywitz, Shaywitz, and Rourke (1996) stated that using a discrepancy score between IQ and achievement to define a learning disability has psychometric concerns because measures of IQ and measures of achievement tap the same cognitive processes that might produce a disability in the first place, implying that there is a great deal of redundancy between IQ and measures of achievement and cognition. Siegel (1989) gave a different reason for hesitancy in using IQ tests, indicating that IQ is not effective in measuring one's achievement ability and concurrently not sufficient in detecting learning disabilities, the overarching reason for this being that measures of IQ are imprecise and unspecific to real-life tasks like reading or math. Flynn (1987) agreed, arguing that problems on IQ tests are "so abstracted from reality that the ability to solve them can diverge over time from the real-world problem-solving ability called intelligence;" he argued that they instead measure some weak correlate to intelligence. Therefore, using IQ as a predictor of math outcome would likely show much overlap with the cognitive constructs being measured, and the extent and the specificity of the redundancy would be difficult to define.

Math Interventions

One of the end goals in understanding the nature of mathematics learning is to optimize ways of teaching and remediating mathematics. For as little that is known about algebra and higher-level math in general, even less is understood about how to teach it or remediate it. Most published interventions are at the elementary school math level (e.g., Fuchs, Powell, Seethaler, Fuchs, Hamlett, Cirino, & Fletcher, 2010; Ma & Kessel, 2003; Poncy, McCallum, & Schmitt, 2010; Xin, Wiles, and Lin, 2008), and there is less empirical data regarding learning algebra and other higher level math. Nonetheless, existing studies introduce intervention structures that are applicable across both developmental and skill-level continuums and may therefore have applicability to higher-level math as well (Fuchs, Fuchs, Powell, Seethaler, Cirino, & Fletcher, 2008).

A common theme in the literature with young students is the teaching of procedural math fact skills. What is known about teaching in elementary math can be translated into higher-level math as well, so its examination is worthwhile. In teaching math fact fluency, the behavioral technique of Cover, Copy, and Compare (CCC), which consists of looking at a math fact and answer, covering it, copying it without looking, and finally comparing the copied response with what the covered math fact was, was more effective than other less direct techniques when employed over the course of a month (Poncy, McCallum, & Schmitt, 2010). Another technique that aims at teaching math fluency is Detect, Practice, Repair (DPR) which incorporates aspects of CCC but also incorporates timed components that build automaticity in math fact retrieval. Students who received DPR instruction each day over the course of four weeks showed significant improvement in math fact fluency (Axtell, McCallum, Mee Bell, & Poncy, 2009). A study by Powell, Fuchs, Fuchs, Cirino, and Fletcher (2009) illustrates the differential effectiveness of math interventions based on the nature of a student's math deficit. Students with math difficulty alone improved more on math fact retrieval after receiving direct fact retrieval training three times a week for fifteen weeks than after receiving computation/estimation training or after receiving no training. On the contrary, students with concurrent math and reading difficulty did not respond differently to any of the interventions.

Procedural interventions tend to be effective for elementary math, possibly because learning basic math facts is an automated process (Axtell et al., 2009) and may not require a great deal of conceptual knowledge. However, for higher-level math, procedural interventions may not be sufficient. Herscovics and Linchevski (1994) suggest that difficulty in algebra may be a result of teaching that does not emphasize learning conceptual strategies. There is a need for conceptual-based teaching strategies in higher level math (Mason, 1989; Sfard & Linchevski, 1994), and students may not receive the full benefits of algebra curricula if procedural math strategies alone are emphasized in higher-level math instruction.

Even at the elementary level, some interventions seek to help students garner a conceptual understanding of math. In fact, Xin, Wiles, and Lin (2008) indicate that teaching basic math at a conceptual level in elementary school may help students to better transition to conceptual learning at higher levels of math. Herscovics and Linchevski (1994) introduced the idea of a cognitive gap between arithmetic and algebra, namely that the lack of conceptual learning in arithmetic makes the abrupt jump to conceptual learning in algebra difficult. An elementary level math intervention called Knowing Math (Ma & Kessel, 2003) involves encouraging the exploration of the mathematical concepts that underlie math procedures. Teachers encourage conversation about strategies, insights, and opinions as well as provide problems which students can use to practice and clarify any confusion. A study of fifth grade students showed that students who received the Knowing Math intervention four times a week for 16 to 20 weeks improved performance significantly more than the control group, which only received regular math instruction from the general education classroom (Ketterlin-Geller, Chard, & Fien, 2008). A case study of three elementary age students with behavioral and emotional disorders featured an intervention emphasizing conceptual understanding of math, math fluency, and problem solving. These students showed signs of improvement in math accuracy and on-task behavior after the intervention's implementation (Alter, Brown, & Pyle, 2011).

As there have been a number of intervention studies targeting elementary math, there have also been studies that examine interventions in higher-level math like algebra. Foegen (2008) reviews some of the types of remedial algebraic instruction. One of these is a cognitive strategy instruction developed by Hutchinson (1993) which involves teaching students how to first represent problems mentally then make a plan for solving problems. Another is creating mnemonics to use as a step-by-step guide for algebraic procedures (Maccini & Hughes, 2000). An additional method uses peer tutoring in the classroom, where one classmate acts as a coach to another, and then they switch roles (Allsopp, 1997). One other method makes use of graphic organizers to help spatially represent simultaneous equation problems (Ives, 2007). Foegen (2008) indicates that the majority of empirical algebra instruction studies, including the aforementioned, have been developed to study remediation for those with math disabilities, not general algebra instruction. While remediation for those with math disabilities is certainly necessary, additional research that seeks to develop robust general algebra instructional practices will help address the *variety* of reasons that a high percentage of students around the United States do poorly in algebra (NAEP, 2011). Additionally, of the studies devoted to examining algebra interventions, few examine college students and/or adults, so there is a need for more research that addresses algebra interventions among adult populations.

An alternative type of intervention, computer algebra systems (CAS), is increasingly being implemented in schools (Taylor, 1995). Anderson, Corbett, Koedinger, and Pelliet (1995) indicate that computer tutoring systems for algebra can be effective and that their effectiveness in remediating algebra skills may be due to a number of principles including that the programs target specific skills, promote abstract understanding of problem-solving knowledge, decrease WM load, and provide immediate feedback for errors. While these kinds of programs have appeal, a study interviewing teachers who used both computer and traditional teaching styles indicated that there might be some downsides as well: teachers reported that computerized instruction was not as effective at encouraging class discussion, that it usually taught only one approach to solving a problem that might have multiple methods of solving, and that it did not inform teachers whether students were thinking about algebraic ideas conceptually or simply at a surface level (Kinney & Kinney, 2002). A large-scale study by the Institute of Educational Sciences (IES) (2009) showed that classrooms which used computerized algebra curriculum did not show any differential improvement in algebra in the first year a teacher used the curriculum. In the second year, classrooms which used the curriculum did slightly better than those that did not use the curriculum, but data was not used from teachers who discontinued use of the curriculum, and the possibility of differential attrition was not controlled. Regardless of the prospective efficiency of the host of algebraic interventions, specific intervention studies need to be done from a variety of perspectives to converge on the most appropriate instructional techniques and principles.

Procedural Skills and Conceptual Knowledge in Algebra

The idea of providing students with a conceptual understanding of math procedures is one that becomes increasingly more important as students transition from elementary math to algebra. Key to algebraic learning is the interplay between procedural and conceptual knowledge. An early study by Thorndike and Upton (1922) theorized that students of algebra struggle because algebra necessitates the ability to “read between the lines” and infer connections that are not always explicitly stated. More recent research expounds upon this idea, suggesting that algebra is cognitively demanding because it requires one to be able to attend to math learning as an object or procedure and concurrently understand and apply the abstraction of that learning (Mason, 1989; Sfard & Linchevski, 1994). In a meta-analysis which included 82 studies and 22,424 students, Rakes, Valentine, McGatha, and Ronau (2010) showed that the effect size for conceptual interventions was nearly double the effect size for procedural interventions. Several studies have shown that lower-level problem solving techniques may actually interfere with learning problem solving techniques that require more conceptual, higher-level understanding (Crowley & Siegler, 1999; McNeil, 2008; McNeil & Alibali, 2005). One instance of this interference effect may play out in the effect learning arithmetic has on learning algebra. Whereas arithmetic emphasizes procedural skills, algebra requires procedural skills in addition to the ability to understand number concepts conceptually, so procedural arithmetic knowledge may interfere with learning conceptual algebraic knowledge (Weaver & Kintsch, 1992). Accordingly, Xin, Wiles, and Lin (2008) touched on the long-term importance of teaching math from a conceptual base even at the elementary level; elementary-age children with or at risk for a math disability who were given a conceptual-based math curriculum showed improvement on measures of equation solving and algebraic expressions.

While the procedural-conceptual dichotomy is a popular one in mathematical learning, Schneider and Stern (2010) suggest that there is still much to understand in order to measure procedural and conceptual knowledge accurately and, subsequently, in order to study their interplay. Rittle-Johnson et al. (2001) demonstrated that students develop conceptual and procedural knowledge in an iterative fashion, i.e., gains in procedural knowledge predicted gains in conceptual knowledge and vice-versa. A study of second grade students found that conceptual instruction improved students' flexibility in choosing procedures for solving math problems (Blöte, Van Der Burg, and Klein, 2001), and because there are multiple ways to solve math problems, choosing the correct procedure can lead to greater knowledge and use of more efficient strategies (Star & Rittle-Johnson, 2008). Therefore, conceptual instruction ought to lead to greater flexibility, which should in turn lead to greater knowledge and efficiency, which should ultimately end with better math performance.

Lee and Hutchison (1998) illustrated the procedural/conceptual paradigm in teaching chemistry stoichiometry principles. Students performed better on a posttest after seeing worked examples and responding to reflection questions that inquired about an underlying strategy than after seeing worked examples and responding to reflection questions that inquired about procedure, although both groups improved from pretest to posttest. Because of the conceptual nature of algebra, it makes sense to teach algebra in a similar manner which guides students to think conceptually – to understand *why* one needs two unique equations to solve for two unknowns, to understand *why* the intersection of two lines is the point at which the solutions to two equations are equivalent or to understand *why* algebraic variables represent real-life entities – not just to memorize a method.

Many studies focus on teaching novel math concepts to students, but there is considerable need for remediation, especially for adults. Croft (2006) and Martin et al. (in preparation) provide evidence that even after taking algebra classes, sometimes at both the high school and college level, a large percentage of college students still are unable to perform critical algebraic procedures. Because of the success of conceptual interventions for algebra when it is being introduced for the first time, it is likely that conceptual interventions would be effective for remediation when algebra is being re-introduced or re-taught to students who have already taken algebra classes. However, more data is needed. For example, it is possible that procedural and conceptual interventions would both be enough to re-activate prior knowledge in students who have already learned the material, and the greater contextualization that a conceptual intervention offers would have no differential effect. Still another possibility may be that neither procedural nor conceptual interventions lead to any change in performance.

Summary

The amount of research devoted to understanding algebra and other advanced math acquisition in adults is less extensive than that focused on arithmetic development in children. A complete understanding of algebra is critical for college students and adults working in a number of different STEM and non-STEM fields. For students at many colleges and universities, algebra is often an obstacle toward upper level coursework, and an inability to pass such a course, or retain the concepts therein, will have negative consequences for students with both STEM and non-STEM majors.

The population for this experiment is a university where many students struggle with algebra, so this is an ideal setting for studying how college students learn algebra. Main components of algebra at the University of Houston (in sequential chapter order) are: graphs and lines, equations and inequalities, functions, polynomial and rational functions, exponentials and logarithms, and simultaneous (systems of) equations (University of Houston Department of Mathematics, 2011). Most undergraduate students at the University of Houston have taken courses at the introductory algebra level or higher in high school, and all students who attended high school in Texas were, in order to graduate, required to pass statewide tests whose content is the same as the content covered in the university's introductory college algebra class (Croft, 2006). Despite this, according to the University of Houston's Office of Institutional Research Program Evaluation by Croft (2006), between 1996 and 2005 the failure rate of introductory college algebra at the University of Houston ranged from 8.4% to 22.2%, coupled with a dropout rate that varied between 14.2% to 22.1%.

The Present Study

Data from the same population as that proposed for the current study (Martin et al., in preparation) showed that performance on items from a particular algebraic measure that involved

2-equation, 2-unknown (2x2) simultaneous equations, was very low, with only 28.57% of students providing the correct answer. Simultaneous equations is one of the more complex skills in introductory algebra at the University of Houston and therefore may show greater variability in performance; for example, it is last in the curricular sequence of the course textbook (University of Houston Department of Mathematics, 2011). Because it is a skill that is relatively more complex and one for which there is evidence for the need for remediation among this population, studying the effects of a "refresher" on simultaneous equations is optimal because it should reduce the presence of possible ceiling performance on pretest and posttest.

The present study will examine the effect of intervention on re-learning simultaneous equations, while also examining the moderating effects of specific cognitive variables and employing reading and arithmetic measures as covariates. The study will employ two interventions, one procedural and one conceptual, which target re-teaching simultaneous equations to college students. Procedural versus conceptual math interventions are particularly relevant because they are analogous to the type of teaching and reflection questions a typical high school or college student might experience – learning to replicate an algebraic algorithm (procedural) versus organizing the learning of algorithms to allow for the abstraction of that procedural knowledge in order to solve diverse types of problems (conceptual). The procedural intervention will encourage students to think concretely about simultaneous equations examples with the goal of understanding *what* is being carried out. On the other hand, the aim behind the conceptual intervention is to encourage students to think through simultaneous equations examples with the mission of figuring out *why* certain procedures are carried out, not just that they *are* carried out. Reflecting what is seen in the literature (e.g., McLean & Hitch, 1999; Tolar et al., 2009; Blair & Razza, 2007; Jiban & Deno, 2007), moderators included in this experiment will include measures of WM, LTM, and visuospatial EF; measures of reading, computational skill, and motivation/anxiety will be used as covariates.

The goals of the study are three-fold. One goal is to examine the roles of WM, LTM, and EF in the efficacy of an algebra intervention. The second goal is to directly compare the effects of a certain procedural and a conceptual intervention, modeled after the instructional paradigm of Lee and Hutchison (1998) referenced above, but with algebra-related worked examples, rather than chemistry-related worked examples. The third goal concerns moderation of the treatment effect by specific cognitive factors, i.e., whether different cognitive factors interact with the results of the procedural or conceptual interventions. Recognizing which type of instruction is most effective for students with different cognitive strengths and weaknesses may be important in crafting math curriculum and lesson plans relevant to student needs.

The present study is unique in many ways that enhance the mathematical literature: (1) its focus is on algebra; (2) it focuses on adult students who have recently taken algebra across a range of mathematical skill; (3) it emphasizes a refresher/re-learning approach that is brief (approximately 15-minute) rather than an intervention that takes several weeks or months; (4) it studies conceptual and procedural differences in algebraic learning; and (5) it studies cognitive correlates of algebraic learning and controls for relevant non-cognitive factors. For these reasons, the study is potentially far-reaching and widely relevant. The implications are significant in that the study seeks to acquire knowledge in several different areas in which research is sparse or weak, making it of interest to those concerned with many different aspects of math learning.

Hypotheses

The first hypothesis is that WM, LTM, and EF are positively correlated with math performance as other studies have shown (Männamaa et al., 2012; Prevatt et al., 2010). The current study expands on previous studies by considering algebra specifically in addition to a broad computational measure.

A second hypothesis is that students randomly assigned to the conceptual condition will perform better on posttest than those in the procedural condition. Both treatment groups will receive similar instruction, and both groups will answer questions about worked examples. The main distinction is that the conceptual group will be asked to answer ten questions about worked examples that force them to think holistically about the process, and the procedural group will be asked to answer thirty questions but that are at a more surface level. It is anticipated that the greater contextualization of the problems in the conceptual treatment group will decrease the necessary cognitive load; in other words, there will be less interference from the memory demands of the task (DeCaro et al., 2010). It is predicted that this difference will be even more marked on transfer-type questions.

A third hypothesis concerns the moderation of the treatment effect by WM, LTM, and EF. It is predicted that the gap between the conceptual and procedural conditions will be largest for participants who are low in WM skills (Huang-Pollock & Karalunas, 2010), LTM skills (Prevatt et al., 2010), and EF skills (Passolunghi & Siegel, 2001) relative to those with other combinations of WM, LTM, and EF (see Figure 1). As per hypotheses 1 and 2, gains on outcome are expected to be greatest with high WM, LTM, or EF skills, and greatest with the conceptual rather than procedural treatment. However, the advantage for the conceptual treatment will be less for high WM, LTM, and EF than for low WM, LTM, and EF. Put another way, the conceptual treatment will compensate to a great extent for low WM, LTM, or EF and only modestly for high WM, LTM, or EF. Although there is literature suggesting that anxiety-relieving interventions helps those with high WM more than those with low WM on math performance (Beilock, 2008; DeCaro et al., 2010), the scope of the present intervention is different; the conceptual intervention focuses on decreasing reliance on WM by encouraging conceptual thinking when attempting to solve novel problems. In other words, being in the conceptual group ought to “make up” for having low WM (Huang-Pollock & Karalunas, 2010). Low LTM is expected to be associated with greater improvement in the conceptual group than in the procedural group because greater LTM is associated with greater math performance, and low anxiety is a protective factor against the consequences of low LTM skills (Prevatt et al., 2010). Participants who are low in LTM skills are predicted to differentially benefit most from the conceptual instruction because the conceptual questions will decrease anxiety levels associated with performing the math tasks. The reason it is predicted that low EF will be associated with greater improvement in the conceptual group than in the procedural group is that greater EF is associated with greater math performance (Passolunghi & Siegel, 2001), and by decreasing cognitive demands in the conceptual group, those low in EF are expected to improve (Bull, Espy, & Wiebe, 2008).

Methods

Participants

Participants were 63 undergraduates (12 male, 51 female) at the University of Houston recruited via the Sona system. Demographics are shown in Table 1. Thirty-one students were randomly assigned to the procedural group and 32 to the conceptual group. Sona is a commercially-based software system designed to manage research administration as well as conduct online surveys. When a student desired to participate, the student used the Sona system to sign up for a time when there was a slot available. The consent document was stored online, and students were able to view the document. Students in certain university psychology classes were given extra credit for participation in psychological research, per agreement with their instructor. Undergraduate students age 25 or younger who had taken (in college or previously) algebra or were currently taking algebra were eligible to participate. The study was approved by the university's Committee for the Protection of Human Subjects (CPHS).

Measures

Algebra. To assess algebra, an experimental pretest and two experimental posttests were used. All three tests feature five problems, one at the top of each page, and participants solve and write answers at the bottom of the page. Participants are informed beforehand that they "will earn points based on the work [they] show, not only [their] final answer." *Pretest* consists of five 2x2 simultaneous equations problems, and participants have up to ten minutes to complete the task. *Posttest I* is analogous to Pretest, with five 2x2 problems of similar difficulty, and participants have ten minutes to complete the task. Problems on Pretest and Posttest I are scored on a scale from 0 to 5, with points given according to a detailed rubric for work shown as well as the final answer. *Posttest II* consists of transfer tasks: two 3x3 problems and three 2x2 problems written in story form. Participants have ten minutes. Problems on Posttest II are scored on a scale from 0 to 7, with points given according to a detailed rubric (see Appendix) for work shown as well as the final answer. Thus, the maximum scores for Pretest or Posttest I are 25, and for Posttest II is 35. See Appendix for examples of Pretest and Posttest I and II problems. Test-retest between Pretest and Posttest I is .61 (though this occurs across intervention). Cronbach's alpha for Pretest, Posttest I, and Posttest II are .85, .75. and .76, respectively.

WM. In the *Digits Back*ward subtest of the *Test of Memory and Language-2* (*TOMAL-2*), participants are verbally given strings of numbers varying from 2 to 9 numbers in length. Numbers range from 1 to 10. After hearing each string of numbers, participants are asked to repeat the string in backward order. The first string is 2 numbers in length, and strings become progressively longer. The experimenter discontinues testing when, after the first 4 items are given, participants correctly answer 3 or fewer numbers correctly on two consecutive items. The total sum of numbers correct is recorded (not the total sum of items). Test-retest reliability is .87 (Reynolds & Voress, 2007).

LTM. The *TOMAL-2 Word Selective Reminding* subtest asks participants to listen to a list of twelve nouns between three and six letters in length, then repeat as many of the items as possible. Items not remembered are repeated by the examiner, and the participant is asked to recall the entire list again. The examiner continues to give feedback about missed items on successive trials until six trials have passed or until the participant successfully recalls all twelve words. The total number of words recalled is recorded; if the participant successfully recalls all twelve words before the sixth trial, credit is given for all remaining trials. Test-retest reliability is .73 (Reynolds & Voress, 2007). In the *TOMAL-2 Word Selective Reminding Delayed* subtest, participants are asked to repeat the twelve-word list after a 20-30 minute delay. Total words recalled are recorded. Test-retest reliability is .47 (Reynolds & Voress, 2007). The measure utilized in analysis is the standard score for immediate recall.

EF. In the *Planning* subtest of the *Woodcock-Johnson III (WJ-III) Tests of Cognitive Abilities*, participants trace over a figure composed of a dotted line, but are not permitted to trace over a line over which they have already traced. The items are scored based on the number of line segments not traced. Internal consistency reliability is .75 (McGrew, Schrank, & Woodcock, 2007). The measure utilized in analysis is the standard score.

Reading. The blue form of the *Word Reading* subtest of the *Wide Range Achievement Test 4 (WRAT-4)* features a list of 55 words of varying difficulty, and participants are asked to pronounce each word. Items are given a score of 1 if they are pronounced correctly and scored 0 if there is any mistake in pronunciation. Internal consistency reliability ranges from .90 to .92 for ages 17-34, and alternate-form delayed-retest reliability is .85 for all adults (Wilkinson & Robertson, 2006). The measure utilized in analysis is the standard score.

Initial Math Skill. The blue form of the *Math Computation* subtest of the WRAT-4 is a 40-item paper-and-pencil math test ranging from single digit addition to fraction manipulation and long division. Items are scored 1 if the answer written is correct and 0 if it is incorrect. Internal consistency reliability ranges from .93 to .94 for ages 17-34, and alternate-form delayed-retest reliability is .88 for all adults (Wilkinson & Robertson, 2006). The measure utilized in analysis is the standard score.

Reflection Questions. The primary purpose of the experimental reflection questions is to foster either procedural or conceptual thinking, depending on the group to which participants are randomly assigned. Secondary to that goal, answers to reflection questions measure algebra aptitude in 2x2 simultaneous equations; scores on *conceptual reflection questions* indicate conceptual aptitude, and scores on *procedural reflection questions* indicate procedural aptitude.

Following the paradigm of Lee and Hutchison (1998) in teaching chemistry stoichiometry problems, the conceptual group is presented with questions that stimulate the formation of a framework for approaching mathematical problems (“why” questions) and the procedural group is presented with questions that support the memorization of rote algebraic techniques (“what” questions). For example, the conceptual group answers questions such as, “In step 2, why did the expert substitute 6-x for y?” and the procedural group answers questions such as, “In step 2, what did the expert substitute for y?”

There are 30 procedural questions and 10 conceptual questions, which reflects the difference in time generally needed to think through the procedural versus conceptual questions.

Procedural questions will be scored as 1 if correct and 0 if incorrect. Conceptual questions will be scored as 3 if complete, 2 if mostly complete, 1 if minimally complete, and 0 if completely incorrect. An example problem and procedural and conceptual questions are presented in the Appendix.

Math Perception. Math perception is assessed using an experimental written questionnaire with 14 items, each assessed on a Likert scale from 1 (Definitely Disagree) to 5 (Definitely Agree). Some items are positively-keyed and some are negatively-keyed. Example items include: "I dread having to do math"; "I would like to learn more math." Negatively-keyed items are reversed in point value when totaling scores. There are three subscales: *Math Anxiety, Perceived Math Importance,* and *Perceived Math Difficulty*. See Appendix for example items. Cronbach's alpha for Math Anxiety is .63, for Perceived Math Importance is .84, and for Perceived Math Difficulty is .77.

Procedure

Participants were given tasks in the following order: Consent; Demographic Questionnaire; Math Anxiety Questionnaire; Previous Math Courses Questionnaire; Pretest; TOMAL-2 Word Selective Reminding; WRAT-4 Math; TOMAL-2 Digits Backward; WRAT-4 Reading; TOMAL-2 Word Selective Reminding Delayed; Reminder Algebra Instruction; Reflection Questions (either conceptual or procedural); WJ-III Test of Cognitive Abilities Planning; Posttest I; Posttest II. The fixed order of administration was a strategic decision 1) in order to eliminate variability due to differences in test administration, 2) in order to maintain a consistent portion of time between Pretest and instruction and between instruction and the posttests, and 3) to allow for the proper amount of time between the TOMAL-2 Word Selective Reminding task and the Word Selective Reminding delayed task. Administering the tasks in reverse order, for example, would have been nonsensical due to the mandatory sequential nature of certain tasks (like the Pretest, posttests, instruction, reflection questions, and TOMAL-2 Word Selective Reminding Delay Task).

For the Reminder Algebra Instruction, a research assistant provided an introduction/recap of the content area, verbally explaining what simultaneous equations are, indicating how to solve simultaneous equations, and writing examples on a marker board. Subjects were presented with two methods of solving such problems procedurally (substitution and elimination). Subjects in the conceptual group received an extra 1-minute verbal explanation of how knowledge about 2x2 simultaneous equations can be extended to simultaneous equations with more than two unknown values and also real-life situations.

For the Reflection Questions, a research assistant placed in front of the participant pages with worked examples that illustrated the solution of 10 different simultaneous equation problems. Underneath each worked example were either conceptual- or procedural-style reflection questions, depending on the group to which they were assigned. In order to maintain equal time for answering reflection questions between the two groups, the procedural group answered 3 questions about each worked example instead of 1 because in pilot testing the conceptual questions took about 3 times as long to answer.

The time of initial instruction was approximately 6 minutes for both groups; the average time of reflection questions was 8.7 minutes (approximately 15 minutes total). Although every effort was made to keep the amount of time for reflection questions similar for both groups, the conceptual group did spend slightly longer on the reflection question portion (9.4 minutes average) than the procedural group (8.0 minutes average). Total participation time varied between about 1 hour and 30 minutes and 1 hour and 45 minutes.

Analyses

The primary analytic approaches include correlation, simultaneous regression, analysis of covariance (ANCOVA), and analysis of variance (ANOVA). Prior to analyses, key variable distributions were examined for their distributional properties (e.g., skew, kurtosis) and outliers. Internal consistency was assessed with Cronbach's alpha. Assumptions for correlation are that data are interval and have a normal sampling distribution (Field & Miles, 2010); normality will be assessed by examining distributional properties. Assumptions for regression include linearity, homoscedasticity, and normality (Osborne & Waters, 2002). Linearity was assessed by examining the regression diagnostics, homoscedasticity was assessed by plotting residuals to make sure that variance of errors are consistent across both levels of the independent variable. Assumptions for ANOVA/ANCOVA include that population error scores will be independently and normally distributed and that the error scores will have an expected value of zero and constant variance (Maxwell & Delaney, 2004). The variance assumption was assessed with Levene's test (Maxwell & Delaney, 2004). Given the preliminary nature of the study, in addition to the inferential statistics described above, effect sizes were also examined when comparing groups. Effect size was calculated using Cohen's (1988) d derived from its original form as follows

In order to control for analyses that either used or did not use covariates, calculation for Cohen's d was based on F rather than directly on the standard error of the samples.

Pretest was part of the design and so was always included in models, and as noted above, was strongly related to posttest performance. Additional variables were also considered as potential covariates: age, sex, time spent on reflection questions, reading ability, math anxiety, perceived math importance, perceived math difficulty, and initial math skill. These covariates were tested to determine if they should be included in fuller models, with the criteria being that they were related to the outcome measures and unrelated to other covariates and the hypothesized predictive variables (Pedhazur, 1997).

Hypotheses

The first hypothesis, that the measures assessing WM, LTM, and EF are predictors of math performance, was tested using bivariate correlation to examine whether each cognitive variable is a significant predictor of posttest math performance. Next, forced entry multiple regression with all three cognitive variables (WM, LTM, and EF) was performed to identify unique contributions of the hypothesized variables. Finally, pretest was added to the multiple regression models, followed by other relevant covariates.

The second hypothesis, that students randomly assigned to the conceptual condition will perform better than those in the procedural condition upon posttest, was tested using ANCOVA. Randomization ensured that Pretest scores were equal for both groups, but the Pretest was used as a covariate to increase power (Knapp & Schafer, 2009). In the context of the treatment effect, we evaluated the pretest variable for assumptions of 1) independence of the covariate and treatment effect and 2) homogeneity of regression slopes (Field & Miles, 2010).

The third hypothesis, that the gap between conceptual and procedural groups on posttest will be greatest for low WM, low LTM, and low EF was assessed by testing for an interaction between each cognitive variable and posttest performance using factorial ANOVA. Because somewhat different expectations are hypothesized, these potential cognitive moderators were examined individually rather than simultaneously.

Power

The study was powered for our primary hypothesis, which concerns the group differences between the experimental groups. There are few existing empirical studies on which to base potential effect sizes. However, in Lee and Hutchison (1998), one of the trials in which participants received either procedural or conceptual-type chemistry reflection questions used 41 subjects, $t(39) = 2.77$, $p < .01$. This corresponds to a Cohen's d effect size of .865. If a similar effect size were obtained in the present study, a sample size of approximately 20 per condition would be needed in order to detect a significant difference between groups with power=.80 and α=.05, suggesting a minimum overall sample size of 40 students. However, in order to maximize impact, 63 students were recruited; at power = .80 and $\alpha = .05$, the minimum detectable effect size would be .70.

Results

Preliminary Results

Descriptive statistics for the predictor and outcome variables are shown in Table 2. There were three cognitive variables (TOMAL-2 Word Selective Reminding, TOMAL-2 Digits Backward, and WJ-III Planning), three measures of anxiety/perception toward math (math anxiety, perceived math importance, and perceived math difficulty) and two academic measures (WRAT-4 Math and WRAT-4 reading). All the variables were normally distributed except WJ-

III Planning, which was skewed to the right and leptokurtotic, and Posttest I, which was skewed to the left and leptokurtotic. However, this skew was exaggerated by the effects of large studentized residuals for both WJ-III Planning and Posttest I. Three individuals in WJ-III Planning and four individuals in Posttest I were found to be excessively influential to the entire sample results, and, therefore, analyses were conducted both with and without these individuals. When removed, the distributions of both of these variables improved substantially, but there was no difference in results when analyses were run with or without them.

At pretest, there were four participants subjects who were at ceiling (a perfect score of 25 out of 25), and nine additional participants at floor (a score of 0 out of 25); follow-up analyses evaluated the effect of these participants on results. Pretest ceiling performers were removed because these participants would be unable to improve from pretest to posttest, and Pretest floor performers were removed because these participants showed no evidence of even partial prior knowledge of the task, which potentially obscures the instruction's status as a "refresher intervention" for these participants. It is not clear whether these participants scored 0 on Pretest because of lack of attention, motivation, or skill. Floor performers were not significantly different from the rest of the sample in any cognitive ability or in reading or math perception ($p \geq .30$), so that their differential performance was restricted to the algebraic task used in this study.

The relationships of covariates to outcomes are shown in Table 3. As noted, initial math ability, time spent on reflection questions, math anxiety, and perceived math importance were significantly related to outcomes, but age, sex, reading ability, and perceived math difficulty were uncorrelated. Therefore, covariates that were significantly related to outcomes were considered in analyses. Other relationships not considered for the analyses, but which may be of interest, are shown in Table 3.

Analyses were run both with and without the relevant covariates for each outcome; however, beyond Pretest, in no case did the substantive results change. Therefore, in the results section, analyses are presented with only Pretest included as a covariate.

Hypothesis 1: WM, LTM, and EF will be positively correlated with math performance. Correlations between all variables are found in Table 3. There were no significant correlations between WM, LTM, and EF and any of the math outcome variables. Correlations ranged from .11 to .20 between WM (TOMAL-2 Digits Backward) and the outcome variables, from .19 to .24 between LTM (TOMAL-2 Word Selective Reminding) and the outcome variables, and from .06 to .17 between EF (WJ-III Planning) and the outcome variables. All of the math outcome variables, including Pretest, Posttest I, Posttest II, and initial math ability were significantly correlated with each other ($p \leq .0001$). When multiple regression was performed with WM, LTM, and EF together and each math outcome variable separately, between 4.8 and 7.8% of the total variance was accounted for by the cognitive measures. The results did not change when additional covariates were included in these regression analyses.

In follow up analyses where floor and ceiling performers were excluded groups stayed relatively even; there were 24 participants in the procedural group and 26 in the conceptual group. Table 4 shows that LTM was now correlated with Posttest I, $r = .36$, $p = .01$, but there were no other significant correlations between LTM, WM, or EF and the different math outcomes.

Hypothesis 2: Students in the conceptual condition will outperform those in the procedural condition at posttest. Table 4 shows a comparison of treatment groups on outcome variables of interest. As expected with the randomized design, there was no difference between treatment groups at Pretest, $F(2,60) = .46$, $p = .50$. To minimize error due to variation on pretest skills within groups, Pretest was used as a covariate in the analysis. There was an overall effect of treatment; i.e., participants scored higher on Posttest I than on Pretest, $F(2,60) = 50.13$, $p < .0001$, $d = 1.78$. There was no significant difference between the conceptual and procedural groups on Posttest I, $F(2,60) < 1$. Effect size was $d = .12$. There was also no significant difference between the conceptual and procedural groups on Posttest II, $F(2,60) < 1$. Effect size was $d = .00$. The results did not change when additional covariates (beyond Pretest) were included in the analysis.

In the follow-up analyses that excluded floor and ceiling performers, there was still an overall effect of treatment from Pretest to Posttest I, $F(2,47) = 47.47$, $p < .0001$, $d = 1.95$, but there remained no difference between the conceptual and procedural groups, $F(2,47) = 3.50$, $p = .068$. Effect size was $d = .53$. There was no significant difference between the conceptual and procedural groups on Posttest II, $F(2,47) < 1$, $p = .40$. Effect size was $d = .24$. When only subjects who were at ceiling on Pretest were excluded, and analyses were run without Pretest as a covariate, there was no difference between groups, $F(2,56) < 1$, $p = .89$.

Hypothesis 3: The gap between the conceptual and procedural conditions will be larger for participants who are low in WM, LTM, or EF relative to those with high WM, LTM, or EF. There were no interactions between any cognitive variable and either Posttest I or Posttest II performance (all $p > .05$). In many ways, this result follows from the first two hypotheses, which noted weak relationships between cognitive variables and posttest performance and the lack of clear difference between the experimental subgroups. The results did not change when additional covariates were included in the analysis. Testing hypothesis 3 without ceiling and floor performers did not alter results; there remained no interactions (all $p \geq .05$).

Discussion

The goal of this study was to evaluate a brief remediation for algebraic simultaneous equations in a conceptual versus procedural experiment, and to understand the influence of individual cognitive differences on experimental effects. While some previous studies assess efficacy of conceptual versus procedural math interventions (e.g., Blöte et al., 2001; Xin et al., 2008) and other studies assess the relationship of cognitive variables to math performance (e.g., Mazzocco & Kover, 2007; Lee et al., 2009; Prevatt et al, 2010), this study is unique in that it assesses the efficacy of algebra interventions under the context of differing cognitive variables.

This study contributes to existing literature in at least three ways. First, by virtue of this study being an experimental study devoted to algebra interventions, it addresses the shortage of research on algebra, particularly the shortage of experiments that compare the effectiveness of different algebraic interventions (Foegen, 2008). Second, participants' performance substantially improved from Pretest to Posttest I on an advanced mathematical skill after an intervention that was only approximately 15 minutes long. Third, although the initial results did not conform to hypotheses, one of the follow-up analyses did reveal an advantage for the conceptual treatment group relative to the procedural group in terms of educationally meaningful effect size ($d = +0.53$).

Was overall treatment effective?

From Pretest to Posttest I, participants improved, which suggests that while there was not a significant difference between treatment groups in the initial analyses, overall, participants scored better on Posttest I than Pretest, which were different forms of the same test (five procedural 2x2 systems of equations problems). Cohen's d effect size is 1.78 which is large. Considering the fact that the intervention was only about 15 minutes total, the intervention was efficient with regard to time, compared to most other intervention studies that intervene over the course of several weeks or months (e.g., Ketterlin-Geller et al., 2008; Hutchinson, 2003; Allsopp, 1997).

It is possible to argue that Posttest I was simply easier than Pretest, or that improvement was due solely to practice effects. However, problems were matched according to difficulty beforehand and split evenly between Pretest and Posttest I, suggesting equivalent difficulty. Also, if practice effects drove the results, then *within* the Pretest and Posttest I, an upward linear increase in scores at the item level would be expected, and there was no such trend (i.e., in the sample as a whole, percentage of participants receiving full credit for each of the five Pretest items was, in order, 52.4%, 61.9%, 50.8%, 28.6%, and 17.5%, and 77.8%, 66.7%, 77.8%, 63.5%, 30.2% received full credit for each of the 5 Posttest I items). If anything, participants tended to do more poorly on later problems. Therefore, it is reasonable to conclude that improvement on the posttest was due to the instruction and/or reflection questions, though differentiation between the type of reflection questions (conceptual or procedural) was not clearly demonstrable in the present case. Whether improvement was due to the instruction, reflection questions, or some combination of the two is unclear.

Did students in the conceptual condition perform better on posttest than those in the procedural condition?

Students in the conceptual condition did not perform significantly better on Posttest I or Posttest II than those in the procedural condition, whether or not covariates were used in the analysis. Research that examines procedural and conceptual interventions tends to favor the use of conceptual interventions (Rakes et al., 2010; Xin et al., 2008, Blöte et al., 2001). Therefore, the results of this study did not concur with existing studies.

It is possible that the study was underpowered. However, the current sample size was powered after an effect found in the literature for a similar type of intervention (i.e., Lee & Hutchison, 1998), and recruitment for the present study surpassed that goal. However, it is still the case that a larger sample may have helped in more clearly establishing intervention effects, particularly given the promising though tentative results of the follow up analyses. Power for detecting a significant difference between posttest performance in the conceptual and procedural conditions was also unexpectedly lowered due to ceiling-level (and floor-level) performance by a substantial portion of subjects. Participants who were at or near ceiling on Pretest could improve their conceptual or procedural efficiency, though this could not be demonstrated with the test utilized here.

Another issue that may have masked a difference between the two groups on posttest is the effect of the reminder instruction itself. The instruction was the same for the two groups, and only the reflection questions differed. Therefore, it is possible that the instruction had so strong an effect that it masked any effect that the reflection questions afforded. The instruction was a refresher; in other words, it was instruction that the participants had already heard before. Because of this, perhaps the instruction itself was enough for the participants to remember how to do the algebra problems, and the reflection questions were unable to contribute to this remembering process above and beyond the effect of the instruction. In a meta-analysis of math remediation studies, Bahr (2008) postulated that “math remediation works for some students,” and “when remediation works, it works extremely well.” One factor Bahr (2008) used to decide whether remediation effectiveness was likely was the depth of remedial need (greater remedial need implying greater remediation effectiveness). Because the depth of remedial need on the present task was not immense for most of the participants (average score on the Pretest was fairly high), perhaps very little remediation was needed for participants to remember the concepts and techniques. Therefore, the instructive part (the part that was consistent for both groups) apart from the reflection questions (the part that differed among groups) may have been sufficient in and of itself to remediate understanding of the algebraic concept, masking any differentiation between the effects of the two different types of reflection questions. Therefore, future research might utilize a control group that does not include reflection questions or does not include instruction.

While the difference was not significant, those in the conceptual condition did score directionally higher at Posttest I than those in the procedural condition. When a portion of the sample was removed due to ceiling and floor effects on Pretest and analyses were re-run, the effect size was moderate favoring the conceptual group, though still not statistically significant ($p < .07$). Therefore, the hypothesized difference between conceptual and procedural methods of intervention may yet hold benefit, and this deserves further study.

Were WM, LTM, and EF positively correlated with algebra performance?

None of the cognitive measures correlated with math performance on Pretest, Posttest I, Posttest II, or on WRAT-4 Math Computation. The only significant relationship identified was in follow-up analyses, showing that LTM with positively related to posttest algebraic performance, which makes particular sense in the present study where participants were required to remember information presented to them several minutes before Posttest I was given. The size of the correlation between LTM and math in the present study ($r=.36$) was similar to findings by Prevatt et al. (2010) ($r=.30$).

It is unlikely that the lack of findings in general was due to the use of a poor outcome measure, as the specific algebra skill measures (Pretest and posttests) did relate well to another more general math test (WRAT-4 Math Computation) which lends validity to them, as others have also shown relationships between algebra and math construed more broadly (Tolar et al., 2009). Further, algebra performance was also related to reading performance, which is in line with previous research that indicates that the ability to read and the ability to do math go hand-in-hand (e.g., Seethaler & Fuchs, 2006). The fact that Posttest II, which included story problems, was correlated with WRAT-4 Reading provides further evidence that the linguistic and textual nature of story problems makes them easier for those with higher reading ability (Powell et al., 2008). Math Anxiety and Perceived Math Importance were also correlated with all of the math variables including Pretest, Posttest I, Posttest II, and WRAT-4 Math. Perceived Math Difficulty was related to Pretest and Posttest II. In other words, participants who were anxious about math, perceived math as unimportant, and perceived math as difficult tended to do worse on math outcome measures. These results are substantiated in previous findings; as other research has shown, an individual's anxiety response to math is inversely associated with how well he or she performs, (Prevatt et al., 2010; Hendel, 1980), and an individual's perception of the importance of math is inversely related to math anxiety (Meece, Wigfield, & Eccles, 1990).

Nonetheless, the lack of relationships between algebra and other cognitive skills such as WM and EF are, in general, in contrast to previous studies (e.g., Mazzocco & Kover, 2007; Lee et al., 2009; Prevatt et al, 2010; Tolar et al., 2009; Bull et al., 2008; Osmon et al., 2006). One possible reason is that the present study used a sample of typical (non-MD) adults, who actually seemed to be better than average at math. Participants in most of the previous studies have examined math performance in participants with MD, not typical-achieving participants. In addition, the mean of WRAT-4 Math performance for the entire sample was at the 79th percentile, well above the overall population average 50th percentile, and higher than expected based on recent previous assessment of the same population (Martin et al., in preparation). Thus, the somewhat restricted range may have mitigated against finding significant relationships.

Additionally, the present study sought to examine a specific algebraic skill in adults, while Tolar et al. (2009) and other algebra studies (Lee et al., 2009; Lee et al., 2011) looked at broad algebraic performance, and other studies (e.g., Swanson & Jerman, 2006; McLean & Hitch, 1999; Hitch & McAuley, 1991; Siegel & Ryan, 1989) examined more basic math skills. Therefore, perhaps the cognitive measures examined in the present study are more related to some aspects of algebra than others. It may also be that different algebra measures (e.g., one that had more difficult problems or more varied types of problems), may have been more sensitive to different cognitive skills (or indeed, to the different intervention groups).

Another issue is that most prior studies look at correlations of cognitive abilities with basic math skills, not with algebra, and use children as participants, not adults. The exception is Tolar et al. (2009) which recruited from undergraduate college students, as did the present study, and found WM to be related to math achievement. However, the size of the relationships in Tolar et al. (2009) (r=.04 to .18) and the present study (r=.11 to .24) were similar, though in the former study the sample size was much larger (N = 195).

Finally, the present study used only one measure of WM, LTM, and EF each, and so it was not possible to comprehensively address each cognitive construct as in a latent variable set up, which would ameliorate measurement issues attributable to any given measure. This is relevant given the variety of measures that can be construed to represent the constructs examined here. For example, EF's relation to some measures of math has been shown using the Tower of London task (Bull et al., 2008) and the Contingency Naming Test (Mazzocco & Kover, 2007), but not visuospatial planning tasks as used in the present study.

Was the gap between the conceptual and procedural conditions largest for participants who were low in WM, LTM, or EF relative to those with high WM, LTM, or EF?

Groups did not perform as predicted, i.e., no interaction was found between the cognitive variables and posttest improvement. While participants in both groups improved from Pretest to posttest, the type of remediation (conceptual or procedural) did not affect participants of varying cognitive abilities in different ways. The same issues with power and sample size due to ceiling performance are applicable. However, the tests for interactions did not approach significance, so it is unlikely that even with a more ideal sample this hypothesis would have been confirmed. The literature focuses more on understanding relationships between math and cognitive areas than how cognitive variables moderate treatment; there were no studies found that explicitly show that cognition moderates treatment effects in a certain direction. Therefore, in examining cognition as a moderator of math remediation, even though significant results were not found one way or the other, the present study explores an area not yet studied.

Conclusion

In sum, there was clearly a significant effect of overall treatment, with participants on the whole improving from pretest to posttest. However, there was no significant difference found between conceptual- and procedural-type algebra remediation, although a follow-up analysis did demonstrate a conceptual advantage with moderate effect size. Besides the fact that LTM was correlated with posttest in a follow-up analysis, WM, LTM, and EF were not found to be significantly correlated with math in most analyses; in contrast, reading and math perception/anxiety appeared to be larger contributors to math performance. Possible reasons for low relationships included sample size, sample shape, examining a narrow algebra topic rather than a broad one, examining adults rather than children, and performance level (typical, rather than MD). There were no significant interactions between WM, LTM, or EF and posttest improvement, which most likely reflects the lack of robust zero-order correlations or differential treatment effects. Despite the lack of robust results directly in line with hypotheses, some promising effects were noted, including substantial overall improvement with only a brief refresher intervention and a moderate effect size advantage of conceptual algebraic interventions over procedural interventions, which warrants future study.

Appendix

Examples of Pretest/Posttest I Questions and Scoring Criteria:

Solve for x and y:
$y = 3x - 2$
$y = -x - 6$

2,2 procedural **(5 pts possible)**
→correct setup of elimination or substitution (more than one way) **(2 pt)**
→correct answer for one variable **(1 pt) | answers are x= -1 and y= -5**
→correctly plugged in first answer into one of the equations – first answer does not have to be correct **(1 pt)**
→correct answer for second variable **(1 pt) | answers are x= -1 and y= -5**

Solve for x and y:
$6x - 3y = 12$
$y = 2x - 4$

2,2 procedural **(5 pts possible)**
→correct setup of elimination or substitution (more than one way) **(2 pt)**
→correct answer for one variable **(1 pt) | answers are x= 1 and y= -2**
→correctly plugged in first answer into one of the equations – first answer does not have to be correct **(1 pt)**
→correct answer for second variable **(1 pt) | answers are x= 1 and y= -2**

Examples of Posttest II Questions and Scoring Criteria:

Bill and Grace have received a total of 64 emails in the past week. If Grace received 5 fewer than twice the number of emails that Bill received, how many emails did they each receive?

2,2 conceptual **(7 pts possible)**
→ correct setup of first eqn **(1 pt) | b+g=64**
→ correct setup of second eqn **(1 pt) | g=2b-5**
→correct setup of elimination or substitution – initial equations do not have to be correct **(2 pt)**
→ correct answer for one variable **(1 pt) | answers are 23 and 41**
→correctly plugged in first answer into one of the equations – first answer does not have to be correct **(1 pt)**
→correct answer for second variable **(1 pt) | answers are 23 and 41**

Solve for x, y, and z:
$3x - 5y + z = 22$
$2x + y = 1$
$x - 3y = 11$

3,3 procedural **(7 pts possible)**
→correct setup of elimination or substitution **(2 pt)**
→correct setup of elimination or substitution a second time **(1 pt)**
→correct answer for one variable **(1 pt) | answers are x=2, y= -3, z=1**
→correctly plugged in first answer into one of the equations – first answer does not have to be correct **(1 pt)**
→correct answer for second variable **(1 pt) | answers are x=2, y= -3, z=1**
→correct answer for third variable **(1 pt) | answers are x=2, y= -3, z=1**

Example of Reflection Question:

The following is a worked example problem:

Find the values for x and y.
Eqn 1: **3x + 4y= 12**
Eqn 2: **18 - 2x= 6y**

1 *Manipulating Equation 1*
3x+4y=12
3(3x+4y=12)
9x+12y=36

2 *Manipulating Equation 2*
18-2x=6y
-2x-6y=-18
2(-2x-6y=-18)
-4x-12y=-36

3 *-4x-12y=-36*
$\underline{9x+12y=36}$
5x=0
x=0

4 *3x+4y=12*
3(0)+4y=12
4y=12
y=3

5 *(0,3)*

Reflection Questions for Procedural Group:

1. In Step 3, which variable was Jeanne trying to eliminate?
2. In Step 1, what does Jeanne multiply the entire equation by?
3. In Step 2, what does Jeanne multiply the entire equation by?

Reflection Question for Conceptual Group:

1. In steps 1 and 2, Jeanne multiplied the equations by different numbers. Why did she choose these numbers?

Math Perception Questionnaire:

1=Definitely Disagree, 2= Somewhat Disagree, 3=Neutral, 4=Somewhat Agree 5=Definitely Agree

I dread having to do math.	1 2 3 4 5
I have forgotten most of the math that I have learned.	1 2 3 4 5
I will just skip a math problem rather than “figure it out” if I don’t immediately know how to answer it.	1 2 3 4 5
I worry about math class more than any other subject.	1 2 3 4 5
I would like to learn more math.	1 2 3 4 5
I would rather take a harder math class if it meant that I learned the material better, even if it meant I would not get as good a grade.	1 2 3 4 5
If I were sitting in a math class, I would be afraid of being called upon.	1 2 3 4 5
It is harder for me to learn math than it is for other people.	1 2 3 4 5
Math is difficult and so it is hard to get motivated to do math related things.	1 2 3 4 5
Math is one of my least favorite subjects.	1 2 3 4 5
Math requires more mental effort than other subjects.	1 2 3 4 5
The thought of doing math makes me nervous.	1 2 3 4 5
There is little point in learning math for my every-day life.	1 2 3 4 5
When I have to do math problems, I “freeze up” and cannot think what to do even though I know I have learned it.	1 2 3 4 5

References

Adams, J. W., & Hitch, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology*, *67(1)*, 21-38.

Allsopp, D. H. (1997). Using classwide peer tutoring to teach beginning algebra problem-solving skills in heterogeneous classrooms. *Remedial and Special Education*, *18(6)*, 367-379.

Alter, P., Brown E. T., Pyle, J. (2011). A strategy-based intervention to improve math word problem-solving skills of students with emotional and behavioral disorders. Education & Treatment of Children (West Virginia University Press), *34(4)*, 535-550.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). *Cognitive tutors: Lessons learned*. *4(2*), 167-207.

Axtell, P. K., McCallum, R. S., Mee Bell, S., & Poncy, B. (2009). Developing math automaticity using a classwide fluency building procedure for middle school students: A preliminary study. *Psychology in the Schools*, *46(6),* 526-538.

Bahr, P. (2008). Does mathematics remediation work?: A comparative analysis of academic attainment among community college students. *Research in Higher Education*, *49(5),* 420-450.

Bahrick, H. & Hall, L. (1991). Lifetime maintenance of high School mathematics content. *Journal of Experimental Psychology: General*, 120(1), 20-33.

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L., & Jacobsen, S. J. (2005). Math learning disorder: incidence in a population-based birth cohort, Rochester, Minn. 1976-82, *Ambulatory Pediatrics, 5(5)*, 281-289.

Beilock, S. L. (2008). Math Performance in Stressful Situations. *Current Directions in Psychological Science, 17(5)*, 339-343.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78(2)*, 647-663.

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.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33(3)*, 205-228.

Cannon, J., & Ginsburg, H. P. (2008). "Doing the math": Maternal beliefs About early mathematics versus language learning. *Early Education & Development, 19(2)*, 238-260.

Center for Education. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Chiappe, P. (2005). How reading research can inform mathematics difficulties: The search for the core deficit. *Journal of Learning Disabilities, 38(4)*, 313-317.

Cirino, P. T. (2011). The interrelationships of mathematical precursors in kindergarten. *Journal of Experimental Child Psychology, 108(4)*, 713-733.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (Second ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooney, J. B., & Swanson, H. L. (1990). Individual differences in memory for mathematical story problems: Memory span and problem perception. *Journal of Educational Psychology*, 82(3), 570-577.

Cowan, N. (2010). Multiple concurrent thoughts: The meaning and developmental neuropsychology of working memory. *Developmental Neuropsychology, 35(5)*, 447-474.

Croft, M. (2006). *An evaluation of the college algebra initiative at the University of Houston*. Report prepared for the Office of Institutional Research, University of Houston, Houston, TX.

Crowley, K., & Siegler, R. S. (1999). Explanation and generalization in young children's strategy learning. *Child Development, 70(2)*, 304-316.

DeCaro, M. S., Rotar, K. E., Kendra, M. S., & Beilock, S. L. (2010). Diagnosing and alleviating the impact of performance pressure on mathematical problem solving. *The Quarterly Journal of Experimental Psychology, 63(8)*, 1619-1630.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11(1)*, 19-23.

Field, A., & Miles, J. (2010). *Discovering statistics using SAS*. London, UK: SAGE Publications Inc.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: what IQ tests really measure. *Psychological Bulletin, 10(2),* 171-191.

Foegen, A. (2008). Algebra progress monitoring and interventions for students with learning disabilities. *Learning Disability Quarterly, 31(2)*, 65-78.

Francis, D. J., Fletcher, J. M., Shaywitz, B. A., Shaywitz, S. E., & Rourke, B. P. (1996). Defining learning and language disabilities: Conceptual and psychometric issues with the use of IQ tests. *Language, Speech, and Hearing Services in Schools*, 27(2), 132-143.

Fuchs, L. S., Fuchs, D., Powell, S. R., Seethaler, P. M., Cirino, P. T., & Fletcher, J. M. (2008). Intensive intervention for students with mathematics disabilities: Seven principles of effective practice. *Learning Disability Quarterly, 31(2),* 79-92.

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., . . . Zumeta, R. O. (2009). Remediating number combination and word problem deficits among students with mathematics difficulties: A randomized control trial. *Journal of Educational Psychology, 101(3),* 561-576.

Geary, D. C. (1993). Mathematical disabilities: cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114(2)*, 345-362.

Geary, D. C., Bailey, D. H., Littlefield, A., Wood, P., Hoard, M. K., & Nugent, L. (2009). First-grade predictors of mathematical learning disability: A latent class trajectory analysis. *Cognitive Development, 24(4)*, 411-429.

Geary, D. C., Hoard, M. K., & Hamson, C. O. (1999). Numerical and arithmetical cognition: patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology, 74(3),* 213-239.

Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2007). Strategy use, long-term memory, and working memory capacity. In D. B. Berch & M.M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 83-105). Baltimore, MD: Paul H. Brookes Publishing Co.

Gersten, R., Clarke, B., & Mazzocco, M. M. (2007). Historical and contemporary perspectives on mathematical learning disabilities. In D. B. Berch & M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilties* (pp. 7-28). Baltimore, MD: Paul H Brooks Publishing Co.

Get Ready to Read. (2011). *About Us.* Retrieved from http://www.getreadytoread.org/about-us

Gierl, M. J., & Bisanz, J. (1995). Anxieties and attitudes related to mathematics in grades 3 and 6. *Journal of Experimental Education, 63(2),* 139-158.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21(1),* 33-46.

Hendel, D. D. (1980). Experiential and affective correlates of math anxiety in adult women. *Psychology of Women Quarterly,* 5(2), 219.

Herriott, S. R., & Dunbar, S. R. (2009). Who takes college algebra? *Primus: Problems, Resources & Issues in Mathematics Undergraduate Studies*, 19(1), 74-87.

Herscovics, N., & Linchevski, L. (1994). A cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics*, *27*, 59-78.

Hitch, G. J., & McAuley, E. (1991). Working memory in children with specific arithmetical learning difficulties. *British Journal of Psychology*, *82(3),* 375-386.

Holmes, J., & Adams, J. (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology*, *26,* 339-366.

Hopko, D. R., Ashcraft, M. H., Gute, J., Ruggiero, K. J., & Lewis, C. (1998). Mathematics anxiety and working memory: Support for the existence of a deficient inhibition mechanism. *Journal of Anxiety Disorders, 12(4),* 343-355.

Huang-Pollock, C. L., & Karalunas, S. L. (2010). Working memory demands impair skill acquisition in children with ADHD. *Journal Of Abnormal Psychology*, 119(1), 174-185.

Hutchinson, N. L. (1993). Effects of cognitive strategy instruction on algebra problem solving of adolescents with learning disabilities. *Learning Disability Quarterly*, *16(1),* 34-63.

Institute of Educational Sciences (2009). *Effectiveness of reading and mathematics software products: Findings from two student cohorts.* Washington, D.C.

Ives, B. (2007). Graphic organizers applied to secondary algebra instruction for students with learning disorders. *Learning Disabilities Research & Practice*, *22(2)*, 110-118.

Jiban, C. L., & Deno, S. L. (2007). Using math and reading curriculum-based measurements to predict state mathematics test performance. *Assessment for Effective Intervention, 32(2)*, 78-89.

Jordan, N. C., & Hanich, L. B. (2000). Mathematical thinking in second-grade children with different forms of LD. *Journal of Learning Disabilities, 33(6)*, 567-578.

Jordan, N. C., & Montani, T. O. (1997). Cognitive arithmetic and problem solving: A comparison of children with specific and general mathematics difficulties. *Journal of Learning Disabilities, 30*, 624-634.

Ketterlin-Geller, L. R. Chard, D. J., & Fien, H. (2008). Making connections in mathematics: conceptual mathematics intervention for low-performing students. *Remedial & Special Education*, *29(1),* 33-45.

Kinney, D.P., & Kinney, L.S. (2002). Instructors' perspectives of instruction in computer-mediated and lecture developmental mathematics classes. In J.L. Higbee, I.M. Duranczyk & D.B. Lundell (Eds.), *Developmental education: Policy and practice* (pp. 127-138). Warrensburg, MO: National Association for Developmental Education.

Knapp, T. R. & Schafer, W. S. (2009). From gain score t to ANCOVA F (and vice versa). *Practical Assessment, Research & Evaluation*, *14(6).*

Lager, C. A. (2006). Types of mathematics-language reading interactions that unnecessarily hinder algebra learning and assessment. *Reading Psychology*, 27(2-3), 165-204.

Lee, A. Y., & Hutchison, L. (1998). Improving learning from examples through reflection. *Journal of Experimental Psychology*: *Applied*, *4(3),* 187-210.

Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology*, *101(2),* 373-387, 387a.

Lee, K., Ng, S. F., Bull, R., Pe, M. L., & Ho, R. H. M. (2011). Are patterns important? An investigation of the relationships between proficiencies in patterns, computation, executive functioning, and algebraic word problems. *Journal of Educational Psychology*, *103(2),* 269-281.

Ma, L. & Kessel, C. (2003). Knowing Mathematics. Houghton Mifflin.

Maccini, P., Hughes C. A. (2000). Effects of a problem-solving strategy on the introductory algebra performance of secondary. *Learning Disabilities Research & Practice*,15(1), 10.

Männamaa, M., Kikas, E., Peets, K., & Palu, A. (2012). Cognitive correlates of math skills in third-grade students. *Educational Psychology*, 32(1), 21-44.

Martin et al., (n.d.) in preparation. University of Houston, Houston, TX.

Mason, J. (1989). Mathematical abstraction as the result of a delicate shift of attention. *For the Learning of Mathematics*, 9(2), 2-8.

Mazzocco, M. M. (2007). Defining and differentiating mathematical learning disabilities and difficulties. In D. B. Berch & M.M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 29-47). Baltimore, MD, US: Paul H Brookes Publishing.

Mazzocco, M. M., & Kover, Sara T. (2007). A longitudinal assessment of executive function skills and their association with math performance. *Child Neuropsychology*, 13(1), 18-45.

Mazzocco, M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia*, 53(1), 218-253.

McGrew, K. S., Schrank, F.A., & Woodcock, R. W. (2007). *Technical manual Woodcock-Johnson III normative update*. Rolling Meadows, IL: Riverside Publishing.

McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology*, 74(3), 240-260.

McNeil, N. M. (2008). Limitations to teaching children 2 + 2 = 4: Typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development*, 79(5), 1524-1537.

McNeil, N. M., & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development*, 76, 883-899.

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology*, 82(1), 60-70.

Miller, S. P., & Mercer, C. D. (1997). Educational aspects of mathematics disabilities. *Journal of Learning Disabilities*, 30(1), 47-56.

National Assessment of Educational Progress. (2011). *The Nation's Report Card.* Retrieved from http://nationsreportcard.gov/math_2011

National Math Advisory Panel. (2008). The final report of the national mathematics advisory panel. Retrieved from http://www2.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, 83(1), 44-59.

Osborne, J. & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*, 8(2).

Osmon, D. C., Smerz, J. M., Braun, M. M., & Plambeck, E. (2006). Processing abilities associated with math skills in adult learning disability. *Journal of Clinical and Experimental Neuropsychology*, 28(1), 84-95.

Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology*, 80(1).

Pedhazur, E. J. (1997). *Multiple regression in behavioral research*. Wadsworth Publishing.

Pickering, S., & Gathercole, S. (2001). *Working memory test battery for children*. London: Psychological Corporation.

Poncy, B. C., McCallum, E., & Schmitt, A. J. (2010). A comparison of behavioral and constructivist interventions for increasing math-fact fluency in a second-grade classroom. *Psychology in the Schools, 47(9)*, 917-930.

Powell, S. R., Fuchs, L. S., Fuchs, D., Cirino, P. T., & Fletcher, J. M. (2008). Do word-problem features differentially affect problem difficulty as a function of students' mathematics difficulty with and without reading difficulty? *Journal of Learning Disabilities, 42(2)*, 99-110.

Powell, S. R., Fuchs, L. S., Fuchs, D., Cirino, P., & Fletcher, J. (2009). Effects of fact retrieval tutoring on third-grade students with math difficulties with and without reading difficulties. *Learning Disabilities Research & Practice, 24(1)*, 1-11.

Prevatt, F., Welles, T. L., Li, H., & Proctor, B. (2010). The contribution of memory and anxiety to the math performance of college students with learning disabilities. *Learning Disabilities Research & Practice, 25(1),* 39-47.

Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of instructional improvement in algebra. *Rreview of eEducational Rresearch, 80(3),* 372-400.

Reynolds, C. R., & Voress, J. K. (2007). *Test of memory and learning*. Austin, TX: Pro-Ed.

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.

Saad, Lydia. (2005). *Math problematic for U.S. teens*. Retrieved from http://www.gallup.com/poll/16360/math-problematic-us-teens.aspx

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology, 96,* 265-282.

Schneider, M., & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology, 46(1),* 178-192.

Seethaler, P. M., & Fuchs, L. S. (2006). The cognitive correlates of computational estimation skill among third-grade students. *Learning Disabilities Research & Practice, 21(4),* 233-243.

Sfard, A., & Linchevski, L.. (1994). The gains and the pitfalls of reification - the case of algebra. *Educational Studies in Mathematics, 26,* 191-228.

Shadish, William R.., Cook, T.D., & Campbell, D.T. (2002). Experimental and quasi-experimental designs for generalized causal inference: Wadsworth Cengage learning.

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60(4)*, 973-980.

Siegel, Linda S. (1989). IQ is irrelevant to the definition of learning disabilities. Journal of Learning Disabilities, *22(8)*, 469-478.

Sikora, D. M., Haley, P., Edwards, J., & Butler, R. W. (2002). Tower of London test performance in children with poor arithmetic skills. *Developmental Neuropsychology*, *21(3)*, 243-254.

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

Swanson, H. L. & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research*, *76*, 249-274.

Swanson, H. L., & Kim, K. (2007). Working memory, short-term memory, and naming speed as predictors of children's mathematical performance. *Intelligence*, 35, 151-168.

Swanson, H. L., Jerman, O., & Zheng, X. (2009). Math disabilities and reading disabilities. *Journal of Psychoeducational Assessment*, *27(3)*, 175-196.

Taylor, M. (1995). Calculators and computer algebra systems: Their use in mathematics examinations. *The Mathematical Gazette*, *79(484)*, 68-83.

Thorndike, E. L., & Upton, C. B. (1922). An experiment in learning an abstract subject. *Journal of Educational Psychology*, *13(6)*, 321-329.

Tolar, T. D., Lederberg, A. R., & Fletcher, J. M. (2009). A structural model of algebra achievement: Computational fluency and spatial visualization as mediators of the effect of working memory on algebra achievement. *Educational Psychology*, *29*, 239-266.

Turner, J. C., Midgley, C., Meyer, D. K., Gheen, M., Anderman, E. M., Kang, Y., & Patrick, H. (2002). The classroom environment and students' reports of avoidance strategies in mathematics: A multimethod study. *Journal of Educational Psychology*, *94(1)*, 88-106.

Tyson, W., Lee, R., Borman, K. M., & Hanson, M. A. (2007). Science, technology, engineering, and mathematics (STEM) pathways: High school science and math coursework and postsecondary degree attainment. *Journal of Education for Students Placed at Risk*, *12(3)*, 243-270.

U.S. Department of Education. (2009). *Reading First*. Retrieved from http://www2.ed.gov/programs/readingfirst/index.html

U.S. Department of Education. (2010). *The Elementary and Secondary Education Act (The No Child Left Behind Act of 2001)*. Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/index.html

U.S. Department of Education. (2012). *Archived: Early Reading First*. Retrieved from http://www2.ed.gov/programs/earlyreading/index.html

University of Houston Department of Mathematics (2011). *Math 1310 college algebra*. Retrieved from http://online.math.uh.edu/Math1310/

Wang, Q., & Pomerantz, E. M. (2009). The motivational landscape of early adolescence in the United States and China: a longitudinal investigation. *Child Development*, *80(4)*, 1272-1287.

Weaver, Charles A., & Kintsch, Walter. (1992). Enhancing students' comprehension of the conceptual structure of algebra word problems. *Journal of Educational Psychology, 84(4),* 419-428.

Wechsler, David. (1997). *Wechsler Adult Intelligence Scale - Third Edition*: The Psychological Corporation.

Wigfield, A., & Meece, J. L. (1988). Math anxiety in elementary and secondary school students. *Journal of Educational Psychology, 80(2),* 210-216.

Wilkinson, Gary S., & Robertson, G.J. (2006). *The Wide Range Achievement Test Administration Manual (Fourth ed.)*. Lutz, FL.: Psychological Assessment Resources, Inc.

Xin, Y. P., Wiles, B., & Lin, Y. (2008). Teaching conceptual model—based word problem story grammar to enhance mathematics problem solving. the journal of special education, *42(3),* 163-178.

Young, C. B., Wu, S. S., & Menon, V. (2012). The neurodevelopmental basis of math anxiety. *Psychological Science, 23(5),* 492-501.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: the role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29(3),* 663-676.

Tables

Table 1. Demographics (N=63)

Variable	Mean (SD) / Frequency
Female	81.0%
Age	21.5 (3.9)
Race	
White	25.4%
Black	23.8%
Asian	25.4%
Hispanic	14.3%
Multiracial	1.6%
Other	9.5%
Right-Handed?	88.9%
Major	
Psychology	54.0%
Other Social Science/Humanities	28.6%
Natural Science/Engineering/Math	17.5%
Year in College	
1	23.0%
2	18.0%
3	26.2%
4	19.7%
5+	13.1%
Transfer Student?	55.6%

Treatment groups do not differ significantly in any of these categories.

Table 2. Descriptive Statistics for Predictor and Outcome Variables

Variable	Category/Scale	N	Mean	Std Dev	Kurtosis	Skewness
Pre/Post Tests						
Pretest	0-25	63	14.19	7.75	-0.72	-0.64
Posttest I	0-25	63	19.32	5.55	2.31	-1.51
Posttest II	0-34	63	20.89	10.27	-1.04	-0.41
Cognitive Tests						
TOMAL-2 WSR	Scaled Score	63	9.83	2.59	0.77	-0.76
TOMAL-2 DB	Scaled Score	62	10.55	2.63	-1.14	0.05
WJ-III Planning	Standard Score	63	107.38	15.93	6.56	2.36
Math Anxiety/Perception						
Anxiety-provoking	1-5	63	2.27	0.71	-0.43	0.32
Unimportant	1-5	63	2.71	0.74	-0.96	-0.06
Difficult	1-5	63	2.46	0.63	0.29	0.24
Other Tests						
WRAT-4Math	Standard Score	63	111.95	13.25	-0.23	0.14
WRAT-4Reading	Standard Score	63	102.14	12.53	0.68	0.23

N = number of participants; Std Dev = standard deviation; WST=Word Selective Reminding; DB=Digits Backward;

Table 3. Correlations between variables (N=63)

	Pre	Post1	Post2	WRAT M	WRAT R	WSR	DB	Plan	Anx	Imp	Diff	Age	SemColl
Pretest	–												
Posttest I	.67**	–											
Posttest II	.51**	.47**	–										
WRAT-4Math	.56**	.51**	.64**	–									
WRAT-4Reading	.12	.19	.23	.39**	–								
TOMAL-2 WSR	.20	.24	.19	.06	.22	–							
TOMAL-2 DB	.20	.18	.11	.24	.23	.24	–						
WJ-III Planning	.17	.10	.06	.22	.02	.13	.42**	–					
Anxiety	-.42**	-.33**	-.31*	-.25*	.00	-.06	-.14	-.22	–				
Importance	-.42**	-.37**	-.41**	-.29*	-.12	-.14	-.16	.00	.63	–			
Difficult	-.31*	-.18	-.25*	-.12	.02	.07	-.07	.00	.72	.46	–		
Age	-.39**	-.19	-.26*	-19	-.13	-.27*	-.09	.05	.27*	.21	.15	–	
Sems. in College	-.28*	-.14	-.24	-.14	.03	-.22	-.05	.01	.01	-.10	.05	.58**	–

*=p<.05, **=p<.01, Pre=Pretest, Post1=Posttest I, Post2=Posttest II, WRATM=WRAT-4 Math, WRATR=WRAT-4 Reading, WSR=TOMAL-2 Word Selective Reminding, DB=TOMAL-2 Digits Backward, Plan=WJ-III Planning, Anx=Math Anxiety, Imp=Perceived Math Importance, Diff=Perceived Math Difficulty, Age=Age of participant, SemColl=Number of semesters participant has been in college. Means and standard deviations can be found in Table 2.

Table 4. Correlations between variables for separate analysis with ceiling and floor performers removed (N=50)

	Pre	Post1	Post2	WRAT M	WRAT R	WSR	DB	Plan	Anx	Imp	Diff	Age	SemColl
Pretest	–												
Posttest I	.49**	–											
Posttest II	.48**	.45**	–										
WRAT-4Math	.47**	.29*	.60**	–									
WRAT-4Reading	.03	.14	.18	.35*	–								
TOMAL-2 WSR	.17	.36*	.22	.07	.22	–							
TOMAL-2 DB	.21	.15	.03	.22	.15	.22	–						
WJ-III Planning	.10	-.08	-.01	.21	-.10	.07	.37**	–					
Anxiety	-.36*	-.38**	-.19	-.19	.08	.02	-.06	-.20	–				
Importance	-.39**	-.37**	-.30*	-.19	-.08	-.13	-.10	.05	.55**	–			
Difficult	-.29*	-.13	-.15	-.05	.08	.16	-.03	.06	.62**	.36*	–		
Age	-.22	-.22	-.25	-.36*	-.09	-.44**	.09	.25	.01	.08	-.14	–	
Sems. in College	-.25	-.25	-.23	-.16	.03	-.25	-.03	.13	-.08	-.14	-.05	.52**	–

*=p<.05, **=p<.01, Pre=Pretest, Post1=Posttest I, Post2=Posttest II, WRATM=WRAT-4 Math, WRATR=WRAT-4 Reading, WSR=TOMAL-2 Word Selective Reminding, DB=TOMAL-2 Digits Backward, Plan=WJ-III Planning, Anx=Math Anxiety, Imp=Perceived Math Importance, Diff=Perceived Math Difficulty, Age=Age of participant, SemColl=Number of semesters participant has been in college. Means and standard deviations can be found in Table 2.

Table 5. Group comparisons on Pretest and outcomes

	All Participants							Ceiling and Floor Performers Removed						
	Conceptual Group		Procedural Group					Conceptual Group		Procedural Group				
	N	Mean (St. Dev.)	N	Mean (St. Dev.)	F	p	d	N	Mean (St. Dev.)	N	Mean (St. Dev.)	F	p	d
Pretest	32	13.53 (7.98)	31	14.87 (7.57)	.46	.50	.17	26	15.69 (5.68)	24	16.08 (4.61)	.07	.79	.08
Posttest I	32	19.25 (6.23)	31	19.39(4.85)	.24	.63	.12	26	21.08 (3.22)	24	19.58(3.87)	3.50	.068	.53
Posttest II	32	20.44(10.62)	31	21.35(10.05)	.00	.99	.00	26	21.96 (10.69)	24	20.13 (10.22)	.72	.40	.24

N = number of participants; Std Dev = standard deviation; d= Cohen's d effect size. For the posttests, F, p, and d were calculated using ANCOVA with Pretest as a covariate.