IN PURSUIT OF KNOWLEDGE: COMPARING SELF-EXPLANATIONS, CONCEPTS, AND PROCEDURES AS PEDAGOGICAL TOOLS

By

Percival G. Matthews

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Psychology

May, 2008

Nashville, Tennessee

Approved:

Professor Bethany Rittle-Johnson

Professor Georgene Troseth

Professor Thomas H. Carr

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

INTRODUCTION

Prompting students to generate explicit explanations of the material they study has emerged as a potentially effective tool for promoting learning and transfer in numerous domains (e.g. Chi, DeLeeuw, Chiu, & LaVancher, 1994). While prompting for such ‘self-explanations’ has been shown to facilitate learning, little is known about how these prompts interact with different types of instruction. Explicating these relations is essential to unlocking the full potential of self-explanation as a tool for supporting learning. Toward this end, the current experiments examined a) whether procedural or conceptual instruction combined with self-explanation prompts differentially affected learning of conceptual and procedural knowledge for children solving math equivalence problems (e.g. $7+3+9=7+\_$); b) whether the type of instruction employed affected the quality of self-explanations generated; and c) and whether self-explanations prompts were effective in promoting learning and transfer over and above conceptual instruction alone.

The Self-Explanation Effect

In their seminal study on the self-explanation effect, Chi, Bassok, Lewis, Reimann, & Glaser (1989) found that, when studying example exercises in a physics text, the best learners spontaneously explained the material to themselves, providing justifications for each action in a solution sequence. Subsequent studies have shown that prompting for such self-explanations can lead to improved learning outcomes in numerous domains including arithmetic (Calin-Jageman & Ratner, 2005; Rittle-Johnson, 2006; Siegler, 2002), geometry (Aleven & Koedinger, 2002; Wong, Lawson & Keeves, 2003), interest calculations (Renkl, Stark, Gruber & Mandel, 1998), LISP programming (Bielaczyc, Pirolli, & Brown, 1995), argumentation (Schworm & Renkl, 2007), Piagetian number conservation (Siegler, 1995), probability calculation (Große & Renkl, 2003), biology text comprehension (Chi, DeLeeuw, Chiu, & LaVancher, 1994), and balancing beam problems (Pine & Messer, 2000). Moreover, these self-explanation effects have been demonstrated across a wide range of age cohorts, from 5-year-old students (Calin-Jageman & Ratner, 2005) to adult bank apprentices (Renkl et al, 1998). Perhaps most impressive is that prompting for self-explanation also promotes transfer in many of these domains, even though participants rarely receive feedback on the quality of their explanations (e.g. Atkinson, Renkl, & Merrill, 2003; Renkl, 1997; Renkl, Stark, Gruber & Mandel, 1998; Wong, Lawson, & Keeves, 2003).

There are, however, substantial differences in the quality of explanations generated among individuals. Importantly, these differences are associated with divergent learning outcomes (Chi et al, 1989; Chi et al, 1994; Pirolli & Recker, 1994; Renkl, 1997). Successful learners tend to give more principle-based explanations, more frequently consider the goals of operators and procedures, and less frequently show illusions of understanding (see Renkl, 2002 for an effective summary). Less successful learners, however, offer fewer explanations, anticipate steps less frequently, examine fewer examples, and tend to focus less on the goals and principles governing operators and procedures (Bielaczyc et. al., 1995; Chi et. al., 1989; Chi et. al., 1994; Pirolli & Recker, 1994; Renkl 1997; Renkl, 1999). Hence, self-explanation prompts are not equally successful across the board at encouraging the types of self-explanations most highly correlated with learning gains. Indeed, a careful review of the literature reveals that prompting learners to self-explain sometimes fails to improve learning at all (Conati & Vanlehn, 2000; Didierjean & Cauzinille Marmeche, 1997; Große & Renkl, 2003; Mwangi & Sweller, 1998; Rittle-Johnson & Russo, 1999).

Thus, while the relation between self-explanation prompts and improved learning has been documented and replicated, much work remains to be done in elucidating the conditions under which such prompts are most effective. Specifically, how can the type of explanations most correlated with improved learning be promoted? One unexplored possibility is that the *type of instruction* preceding self-explanation prompts may influence subsequent explanation quality and learning. Although method of instruction has been varied between experiments, the type of instruction used *within* an experiment has rarely been manipulated. In this study, we contrast the effects of conceptual and procedural instruction on self-explanation quality and learning.

Which Type of Instruction?

Debate over the comparative merits of procedural and conceptual instruction has a rich history spanning the 20^{th} century (see Baroody & Dowker, 2003 for an overview), yet the relations between the types of instruction employed and the types of mathematical understandings generated remain largely unresolved. As pedagogical approaches generally involve some level of tradeoff between the two, it is of both theoretical and practical import that we gain some insight into the nature of the knowledge promoted by each. Does instruction focusing on procedures primarily build procedural knowledge, or does it effectively promote conceptual knowledge as well? Likewise, what types of knowledge does instruction on concepts promote? In line with our present concern for getting the most out of self-explanations, we add another question: which type of instruction best supports the types of explanations associated with the best learning gains?

Part of the difficulty with delimiting the relations between type of instruction and type of knowledge lies with problems inherent to distinguishing between the types of knowledge themselves. Procedural and conceptual knowledge lie on a continuum and are not always easily separated (Rittle-Johnson, Siegler & Alibali, 2001; Star, 2005).

With this in mind, we offer some operational definitions to help clarify the constructs central to our investigation. In accord with previous studies on the topic, we define conceptual knowledge as explicit or implicit *knowledge of the principles* that govern a domain and their interrelations. In contrast, we define procedural knowledge as the *ability to execute action sequences* to solve problems (see Baroody, Feil & Johnson, 2007; Greeno, Riley, & Gelman, 1984; Hiebert & LeFevre, 1986; Rittle-Johnson et al., 2001: Star, 2005). In the current study, we assessed conceptual knowledge by a variety of measures, both explicit – such as asking students what the equal sign means – and implicit – such as asking students whether or not certain mathematical expressions are acceptable. We assessed two components of procedural knowledge: *procedural learning* was assessed as accuracy on math equivalence problems of the same format as those used in the intervention, and *procedural transfer* was assessed as accuracy on math equivalence problems with certain alterations to operators or to the positioning of the unknown quantity. The essential difference between the two kinds of problems is that transfer problems demand that the learned procedure be adapted to novel problem features. Our definitions for type of instruction follow directly from our definitions of these respective types of knowledge. Specifically, we define *conceptual instruction* as instruction that focuses on domain principles and *procedural instruction* as instruction that focuses on step-by-step problem solving procedures.

Several classroom researchers have argued that, compared to procedural instruction, conceptual instruction supports more general and flexible knowledge gains . Hiebert and Wearne's (1996) study of place value and multi-digit arithmetic is one widely cited case. This six-week study measured procedural knowledge as the ability to correctly perform addition and subtraction procedures, and measured conceptual knowledge as students' understanding of place value and its connection with written numerical calculations. For instance, one measure of conceptual understanding required children to use color-coded chips to demonstrate the same computation they completed using paper and pencil. Hiebert and Wearne's conceptually based 'alternative' instruction focused on quantifying sets of objects by grouping by tens, analyzing different forms of representing quantities, and then building addition and subtraction procedures from the ideas underlying the base ten system. Children did receive a limited amount of procedural instruction in this condition. Their procedurally based 'conventional' instruction, in contrast, focused on the standard algorithms for addition and subtraction as suggested by the textbook. The authors found that procedural instruction could quickly move students' procedural knowledge ahead of their conceptual knowledge, whereas conceptual instruction improved both procedural and conceptual knowledge simultaneously. Others also have found evidence that, compared to procedural instruction, conceptual instruction leads to greater conceptual knowledge and to comparable procedural knowledge

(Bednarz & Janvier, 1988; Blöt, Van der Burg, & Klein 2001; Cobb, Wood, Yackel, Nicholls, Wheatley, Trigatti, & Perlwitz, 1991; Fuson & Briars, 1990; Hiebert & Grouws, 2007; Kamii & Dominick, 1987).

Randomized experimental studies have provided some corroboration of these classroom findings. These studies have used math equivalence problems as the target task. In one early precursor to the current experiment, Perry (1991) employed a conceptual instruction in which children were told the goal of the problem but were given no instruction on procedures to achieve the goal. The procedural instruction taught students a procedure for solving the problems. Procedural knowledge was measured as the ability to generate correct answers to math equivalence problems, both with familiar and with novel problem features. Providing students with conceptual instruction led many children to generate accurate solution procedures that they could appropriately adapt to solve transfer problems. By contrast, procedural instruction improved performance on problems specifically targeted by instruction but was less effective in promoting procedural transfer. Rittle-Johnson and Alibali (1999) found similar results, as well as that procedural instruction was less effective than conceptual instruction at promoting conceptual knowledge. Interestingly, Perry (1991) also found that procedural instruction could actually *impede* learning, as students who received hybrid instruction on both concepts and procedures performed worse on procedural transfer items than those who received instruction on concepts alone.

These findings notwithstanding, we should be careful not to conclude prematurely that conceptual instruction is always more effective than procedural instruction in promoting conceptual understanding and procedural transfer. In particular, we should note that the above findings tend not to argue that procedural instruction is completely ineffective – they instead they argue that it is not as effective as conceptual instruction. Indeed, recent experiments have explicitly shown that either procedural instruction or procedural practice in the absence of instruction can promote both procedural and conceptual knowledge of math equivalence and decimals (Rittle-Johnson & Alibali, 1999; Rittle-Johnson et. al., 2001). The question, then, is whether or not there may be conditions under which procedural instruction may prove equally effective or even superior to conceptual instruction. It may very well be that the procedural instruction manipulations discussed above were not well suited to maximize the effects of procedural instruction. It may also be that design limitations did not allow for construction of appropriate comparison groups for analyzing the effects of procedural versus conceptual instruction.

The classroom studies considered above often failed to fully isolate the type of instruction employed, raising the question of whether appropriate groups were used for comparison. What some referred to as conceptual instruction often included elements of procedural instruction (see Hiebert & Wearne, 1996; Blöt, Van der Burg, & Klein 2001). Thus the comparison is not between conceptual instruction and procedural instruction, but instead between some hybrid of the two and procedural instruction alone. Problems also exist in the procedural instruction manipulations of the experimental studies considered above. These studies have either offered few examples, little opportunity for feedback, or not prompts for reflection, all of which may be important for establishing effects of procedural instruction (see Perry, 1991; Rittle-Johnson & Alibali, 1999; Rittle-Johnson et. al., 2001).

All told, these studies may not have allowed for the proper acquisition of robust procedures that are hypothesized to reduce cognitive demand and increase problem solving efficiency (see Anderson, 1993; Kotovsky, Hayes & Simon,1985; Proctor & Dutta, 1995 Sweller, 1998; van Merrienboer, Jelsma & Paas, 1992). The procedural instruction intervention we employ below offers both instruction on a procedure *and* several opportunities for practice using that procedure in addition to offering feedback on participant performance. Moreover, the current study incorporates self-explanation prompts for reflection – discussed below – which may further boost the effects of procedural instruction. Thus, it may result in the swifter acquisition of robust procedures than the manipulations in the investigations discussed above boosting the efficacy of procedural instruction.

Instruction and Self-Explanation

Self-explanation prompts add a new dimension to consider when choosing between procedural and conceptual instruction. The effects of a given type of instruction might be augmented or weakened when used in combination with self-explanation prompts. Likewise, the effects of self-explanation prompts might vary in response to the type of instruction used prior to prompting.

Perhaps conceptual instruction can boost the benefits of self-explanation prompts by directly augmenting knowledge of domain principles and directing attention to conceptual structure. Similarly, self-explanation prompts may help students further fill in mental models by promoting inferences that can be drawn from knowledge provided by conceptual instruction. Alternatively, it may be that procedural instruction frees up cognitive resources that can be dedicated to generating more effective self-explanations when compared to conceptual instruction. To date, the comparison remains unexamined.

The Current Experiments

The current experiments investigated the relations between type of instruction, self-explanation prompts, and the types of self-explanations and knowledge that are promoted. We used math equivalence problems of the type $7 + 3 + 9 = 7 + \_$ as the primary task. These problems pose a relatively high degree of difficulty for elementary school children (Alibali, 1999; Perry, Church, & Goldin-Meadow, 1988; Rittle-Johnson, 2006). Importantly, these problems tap children's understanding of equality, which is a fundamental concept in arithmetic and algebra (Kieran,1981; Knuth, Alibali, McNeil, Weinberg & Stephens, 2005; McNeil & Alibali, 2005). Because equality is such a central concept in mathematics, the present tasks offer a potentially fruitful field for exploration of the relations between conceptual and procedural knowledge in mathematical thinking more generally.

Prior research has shown that self-explanation prompts can improve procedural learning and transfer on math equivalence problems (Rittle-Johnson, 2006; Siegler, 2002). No previous study, however, has investigated the comparative merits of procedural versus conceptual instruction in conjunction with self-explanation prompts. Unpacking these relations is essential both to discovering the mechanisms by which self-explanations may work and to optimizing pedagogical programs based upon them.

The goals of the present study were threefold. First, because of the previously established relation between quality of self-explanation and learning outcomes, we wanted to evaluate the relations between type of instruction and the quality of children's subsequent self-explanations. Second, we wanted to evaluate the relations between type of instruction and children's evolving conceptual and procedural knowledge of mathematical equivalence. Finally, we wanted to determine whether self-explanation prompts used in conjunction with conceptual instruction improve learning over and above conceptual instruction alone when controlling time on task. Specifically, Experiment 1 examined the comparative effects of conceptual and procedural instruction when all children were prompted to self-explain. Experiment 2 examined the effects of self-explanation prompts when all children received conceptual instruction.

EXPERIMENT 1

We expected conceptual instruction to promote superior gains in conceptual knowledge and procedural transfer. We predicted that both groups, however, would show comparable gains for procedural learning problems. These expectations are largely in accord with the findings from Rittle-Johnson and Alibali (1999) and Perry (1991).

We expected behavior during the intervention to vary in accord with these outcomes. First, in light of the limited-resources account sketched above, we anticipated that the procedurally instructed group would show more rapid initial improvement in accuracy. However, by posttest, we expected that the conceptually instructed group would catch up given the opportunities for problem solving and feedback during the intervention. Second, we hypothesized that this rate of improvement would be reflected in children's strategy use. Procedurally instructed students were expected to require less problem solving search, quickly adopting the instructed procedure and using that procedure consistently whereas conceptually instructed students were expected to need time to invent varied procedures during the intervention. Third, we anticipated that conceptually instructed students' self-explanations would be more conceptual in nature than the procedurally instructed students'. In contrast, we predicted that procedurally instructed students would focus less on domain properties in their explanations and more on the repetition or explication of the single procedure specifically taught during the intervention. These differences in explanation quality were expected to predict performance at posttest, with more conceptual explanations predicting higher performance on all three outcome measures.

Method

Participants

Consent was obtained from 121 second- through fifth-grade children from an urban parochial school serving a middle-class, predominantly Caucasian population. A pretest was given to identify children who could not already correctly solve half of the math equivalence problems targeted for intervention. The final sample consisted of 40 children: 14 second-graders (9 girls), 8 third-graders (4 girls), 5 fourth-graders (3 girls), and 13 fifth-graders (6 girls). Their average age was 9.6 years (range 7.5-11.8). Children participated in the spring semester.

Design

Children completed a pretest, intervention, immediate posttest, and a two-week retention test. All were randomly assigned to either the procedural instruction (n=21) or the conceptual instruction condition (n=19). Children from each grade were evenly distributed across the two conditions. During the intervention, children first received instruction and then practiced solving six mathematical equivalence problems. All children received accuracy feedback and were prompted to self-explain on the practice problems. Details of the protocol are discussed in the procedure section below.

Assessments

Identical assessments of conceptual and procedural knowledge were administered at pretest, immediate posttest, and retention test. There were two procedural learning problems (i.e. $7+6+4=7+\_$; $4+5+8=\_+8$). There were also six procedural transfer problems that either a) had no repeated addend on the right side of the equation (i.e. $6+3+5=8+\_$; $5+7+3=\_+9$), b) had the blank on the left side of the equation (i.e. $\_+9=8+5+9$; $8+\_=8+6+4$), or c) included subtraction (i.e. $8+5-3=8+\_$; $6-4+3=\_+3$). At posttest, the learning problem format was familiar, and children could solve them using step-by-step solution procedures learned during the intervention. By contrast, the transfer problem formats remained unfamiliar to the children at posttest, so had to be solved by applying or adapting procedures learned during the intervention – a standard approach for measuring transfer (e.g., Atkinson et al., 2003; Chen & Klahr, 1999). Children were encouraged to show their calculations when solving the problems.

The five items on the conceptual knowledge assessment are shown in Table 1. The items assessed children's knowledge of two key concepts of equivalence problems: (a) the meaning of the equal sign as a relational symbol, and (b) the structure of equations, including the idea that there are two sides to an equation. All items were adapted from Rittle-Johnson (2006) and Rittle-Johnson & Alibali (1999) and were designed to measure both explicit and implicit conceptual knowledge. Each item was

Table 1 - Conceptual Knowledge Assessment Items

Concept	Item	Scoring
Meaning of Equal Sign		
	Define Equal Sign.	1 Point if defined relationally (e.g. 'Two amounts are the same" "Equivalent to" "Same on both sides" "The numbers on each side are balanced")
	Rate definitions of equal sign: rate 4 definitions as "always, sometimes, or never true".	1 Point if student rated the statement, "The equal sign means two amounts are the same," as, "always true".
Structure of Equations		
	Correct encoding: Reproduce equivalence problems, one at a time, from memory after a 5 s delay. Total of four problems.	1 Point if students put numerals, operators, equal sign and blank in correct respective positions for all 4 problems.
	Recognize correct use of equal sign in multiple contexts: indicate whether 8 equations such as 8 = 2+6 or 3+2=7-2 make sense.	1 Point if >75% correct.
Meaning of Equal Sign and Structure of Equation	a) Record the two separate sides of the equation 4+3 = 5+2.	1 Point for part a) if 4+3 and 5+2 are each identified as separate sides of the equation.
	b) State the meaning of the equal sign in *this* problem.	1 Point if defined relationally, as above.

scored as 0 or 1 point for a possible total of 6 points (see Table 1 for scoring criteria), and scores were converted to percentages.

For procedural knowledge items, we coded the procedure each child employed for each problem based on his or her answers as well as written calculations on the assessments or verbal reports given during the intervention (see Table 2). Accuracy scores were calculated based on the percentage of problems children solved using a correct procedure, regardless of whether they made arithmetic errors.

Procedure

Children completed the written pretest in a 30-minute session in their classrooms. Within one week of the pretest, each participant completed a one-on-one intervention and immediate posttest in one session lasting approximately 45 minutes. This intervention session was conducted by the author in a quiet room at the school. The retention test was administered approximately two weeks later in a group session lasting no longer than 30 minutes.

Per Rittle-Johnson (2006), all intervention problems were standard mathematical equivalence problems with a repeated addend on the two sides of the equation, and they varied in the position of the blank after the equal sign (i.e., $4 + 9 + 6 = 4 + \_$ and $3 + 4 + 8 = \_ + 8$, which are referred to as standard A+ and +C problems, respectively).

At the beginning of the intervention, children in the procedural instruction condition were taught an add-subtract procedure using a total of five example problems. They were first instructed on two standard A+ problems. The experimenter often prompted students with questions to ensure that they were attending to and understanding the instruction. For instance, for the problem 3 + 4 + 2 = 3 + _, the experimenter said:

> This is what you can do: you can add the 3 and the 4 and the 2 together on the first side of the equal sign [He then drew a circle around the 3+4+2], and then subtract the 3 that's over here [He underlined the 3], and that amount goes in the blank. So, for this problem, what is 3 + 4 + 2? [Waited for student response] Right, 9, and 9 minus 3 is what? [Waited for student response] Great, so our answer is 6.

After receiving instruction on the two standard A+ problems, students received similar instruction on two +C questions problems and a final A+ problem.

Children in the conceptual instruction condition were taught about the relational function of the equal sign, also using five examples. First children were asked to define the equal sign. They were then given an explicit definition for the meaning of the equal sign, using a number sentence as an example. Specifically, they were shown the number sentence 3 + 4 = 3 + 4, and the experimenter said:

> There are two sides to this problem, one on the left side of the equal sign [He made a sweeping gesture under the left side] and one on the right side of the equal sign [sweeping gesture under the right side]. The first side is 3+4 [sweeping gesture]. The second side is 3 + 4 [sweeping gesture]. What the equal sign [pointing] means is that the things on both sides of the equal sign are equal or the same [sweeping hand back and forth].

Students were shown four other number sentences of various sorts (i.e. $4+4=3+5$; $3+4=\_$; $2+3 \bigcirc 3+6$[1]; $5+4+3=5+\_$) and reminded of what the equal sign meant in each case. This brought the total number of examples to five in order to parallel the number of problems encountered in the procedural instruction condition. No solution procedures were ever explicitly discussed in the conceptual instruction session. As in the procedural condition, the experimenter often prompted students with questions

[1] For this item, students were asked, "would it make sense to write an equal sign here [in the circle]?"

to ensure that they were attending to and understanding the instruction. Instruction took approximately six minutes in either condition.

The remainder of the intervention session was the same for both conditions. Practice problems were six standard mathematical equivalence problems with a repeated addend on both sides of the equation. Problems were presented on a laptop and alternated between A+ and +C problems so that students had experience with problems in which the position of the blank varied. For each of the problems, all children solved the problem, reported how they solved the problem, and received accuracy feedback. Children were then prompted to self-explain. The self-explanation prompt was the same as the one used in Rittle-Johnson (2006), which was adapted from Siegler, (2002). Children saw a screen with the answers that two children at another school had purportedly given: one correct

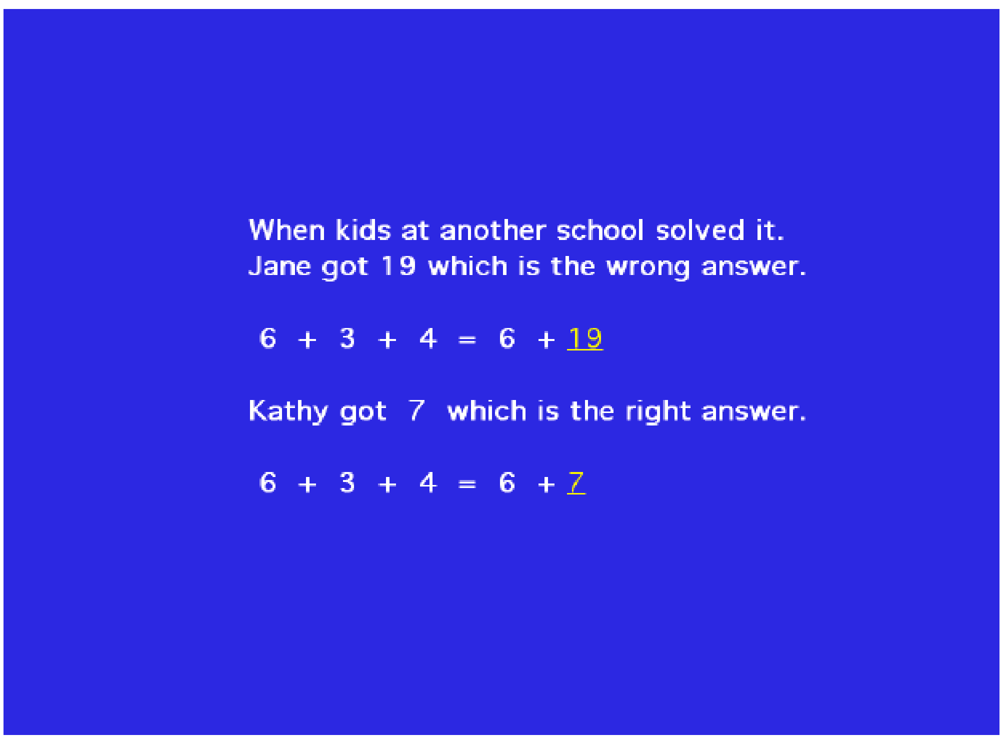


Figure 1. Screen shot of the additional screen for the self-explanation condition.

and one incorrect, as shown in Figure 1. The experimenter then asked the participants both how the other children got their answers and why each answer was correct or incorrect. We asked both questions to highlight for children the distinction between how a procedure is employed and why it is correct or incorrect.

Answers to the 'why' questions were counted as self-explanations. We had children explain the correct and incorrect answers of others because previous work has shown that self-explanation works best when participants are asked to explain correct reasoning instead of their own, sometimes incorrect, reasoning (e.g. Calin-Jageman & Ratner, 2005) and when they are asked to explain both correct and incorrect reasons (Siegler, 2002).

Students' self-explanations of why solutions were correct and incorrect during the intervention were coded. *Procedural explanations* explicitly referenced specific solution steps with no other rationale (e.g. "You would always add those two together first and then you would have subtracted 22 by 6"), *conceptual explanations* referred to the need to make the two sides of an equation equal (e.g. "Because it makes it equal on both sides"), and *other explanations* offered vague responses, nonsense responses, or non-responses (e.g. "That's what the problem tells you to do").

The intervention was audiotaped and videotaped. Total time spent on the practice problems was similar across conceptual, (M = 15.72 min, SD = 3.89 min) and procedural instruction conditions (M = 14.86 min, SD = 3.29 min, $t(38)$ = .76, p = .45). Immediately following the intervention, children completed a paper and pencil posttest administered individually by the experimenter in the same room. Approximately two weeks later, students completed a delayed retention test as groups in their classrooms.

Independent raters coded 20% of participants' procedure use across all phases of the study and their *why* explanations during the intervention. Inter-rater agreement ranged from 81% for self-explanation quality to 90% for procedure use during the intervention.

Treatment of Missing Data.

Three participants (8% of the sample) were absent from school on the day of the retention test (2 in the procedural instruction condition and 1 in the conceptual instruction condition). The absent participants did not differ significantly from those who were present on the pretest measures. To deal with this missing data, an imputation technique was used to approximate the missing accuracy scores on the retention test (Harrell, 2001). Imputation leads to more precise and unbiased conclusions than casewise deletion (Peugh & Enders, 2004; Schafer & Graham, 2002), and simulation studies have found that using Maximum Likelihood (ML) Imputation when data is missing at random leads to the same conclusions as when there is no missing data (Graham, Hofer, & MacKinnon, 1996; Schafer & Graham, 2002).

As the children had no knowledge of the date of the delayed posttest, these data could be considered as missing at random (confirmed by Little's MCAR test: $\chi^2(26) = .58$, $p > .90$). As recommended by Schafer & Graham (2002), we used the EM algorithm for ML estimation via the missing value analysis module of SPSS. The students' missing scores were estimated from all non-missing values that were included in the analyses presented below. Comparison of effect-sizes for the condition manipulation when students with incomplete data were deleted, rather than imputing their missing scores, indicated that the ML estimates had minimal influence on effect-size estimates; imputed data led to effect-sizes that were quite similar to those observed with a case-wise deletion approach (i.e. the change in η^2 was <.02 for all significant variables). There were no substantive difference between analyses conducted with case-wise deletion and imputation.

Results and Discussion

First, we summarize the participants' knowledge base at pretest. This summary is followed by comparisons of children's behavior during the intervention, including their accuracy, procedure use and self-explanation quality. Finally, we report on the variables that affect posttest and retention performance. Effect sizes are reported as partial eta squared (η^2) values.

Pretest

Children who were included in the study had little knowledge of correct procedures for solving mathematical equivalence problems at pretest. Most (62%) did not solve any of the four pretest problems correctly, 15% solved only one problem correctly, and 23% solved two problems correctly. At pretest, children typically added all four numbers or added the three numbers before the equal sign (see Table 2), and there was no significant difference in accuracy between the two conditions, $F(1, 38) = 2.58$, $p = .12$, $\eta^2 = .06$.

Children began the study with some conceptual knowledge of mathematical equivalence (*M* = 38%, *SD* = 21%). Although children were randomly assigned to condition, there was a difference between groups in conceptual knowledge at pretest, with children in the conceptual condition (*M* = 46%, *SD* = 18%) scoring higher than those in the procedural instruction condition (*M* = 32%, *SD* = 22%), $F(1, 38) = 4.73$, $p = .04$, $\eta^2 = .11$. To help control for these differences, pretest knowledge was included as a covariate in all subsequent models.

Intervention

We expected the two conditions to differ in their accuracy, procedure use and self-explanation quality during the intervention. To evaluate this, a series of ANCOVAs

Procedure	Sample explanation	Conceptual Condition				Procedural Condition			
		Pre	Intervention	Post	Retention	Pre	Intervention	Post	Retention
Incorrect Procedures									
Add all	I added the 8, the 8, the 7 and the 3	21	2	3	6	21	0	0	13
Add to equal sign	8 plus 7 equals 15, plus 3 is 18	22	2	1	5	32	0	1	6
Don't Know	I don't know	16	2	5	2	11	0	8	4
Other	Other incorrect – I used 8 plus 8 and then 3	10	9	5	10	19	1	9	16
Used *Any* Incorrect Procedure		***69***	***14***	***14***	***22***	***84***	***1***	***17***	***39***
Correct Procedures									
Equalize	I added 8 plus 7 plus 3 and I got 18 and 8 plus 10 is 18	21	50	55	48	8	2	11	11
Add-subtract	Add/Subtract -I did 8 plus 7 equals 15 plus 3 equals 18 and then 18 minus 8 equals 10	2	17	15	15	2	97	64	41
Grouping	I took out the 8's and I added 7 plus 3	7	13	9	8	0	0	0	0
Ambiguous	8 divided by 8 is 0 and 7 plus 3 is 10	2	6	7	7	7	0	8	9
Used *Any* Correct Procedure		***30***	***86***	***86***	***78***	***16***	***99***	***83***	***61***

Table 2 – Procedure Use in Experiment 1was conducted with type of instruction as a between-subject factor. Conceptual and procedural knowledge pretest scores as well as grade level were included in all analyses as covariates to control for prior knowledge differences. Preliminary analyses indicated that student's grade level never interacted with condition, so this interaction term was not included in the final models.

Accuracy. Procedural accuracy during the intervention was higher for the procedural instruction group than for the conceptual instruction group, $F(1, 35) = 5.22$, $p = .03$, $\eta^2 = .13$. There was also an effect for prior procedural knowledge, as children with higher procedural knowledge pretest scores were more accurate, $F(1,35) = 4.43$, $p = .04$, $\eta^2 = .11$. Prior conceptual knowledge, however, did not influence performance.

Procedure use. As expected, type of instruction also influenced both what procedures children used and how many different procedures they used. Children in the procedural instruction condition adopted the add-subtract procedure the vast majority of the times (see Table 2). Only 3 of 21 children in the procedural instruction group used an identifiably correct procedure other than the add-subtract procedure, and one child was responsible for more than half of all trials solved by a different method. A one way ANCOVA with the frequency of Add-Subtract use as the dependent variable and condition as the independent variable verified that students in the procedural condition were far more likely to use the add-subtract procedure than those in the conceptual condition, $F(1, 35) = 36.64$, $p < .01$, $\eta^2 = .51$.

Children in the conceptual instruction condition, in contrast, employed a number of strategies. They were more than four times as likely to use multiple correct procedures (42% and 10% of children in the conceptual and procedural conditions, respectively, χ^2

(1,40) = 5.65, p =.02). Although they used more correct procedures, students in the conceptual group showed a strong preference for the equalizer strategy and were far more likely than those in the procedural condition to use this strategy, $F(1, 39) = 16.10$, $p < .01$, $\eta^2 = .30$. Altogether, these results support our hypothesis that provision of a robust procedure decreased problem-solving search for the procedurally instructed group, leading to rapid adoption of the instructed procedure. As a consequence, it also led the procedurally instructed group to use fewer different correct procedures overall.

Explanation quality. There was a stark contrast in explanations offered in response to the *why* questions by condition. Children who were given conceptual instruction provided a conceptual rationale on over half of all explanations (M = .54, SD = .34, see Figure 2), whereas children in the procedural instruction condition rarely did so

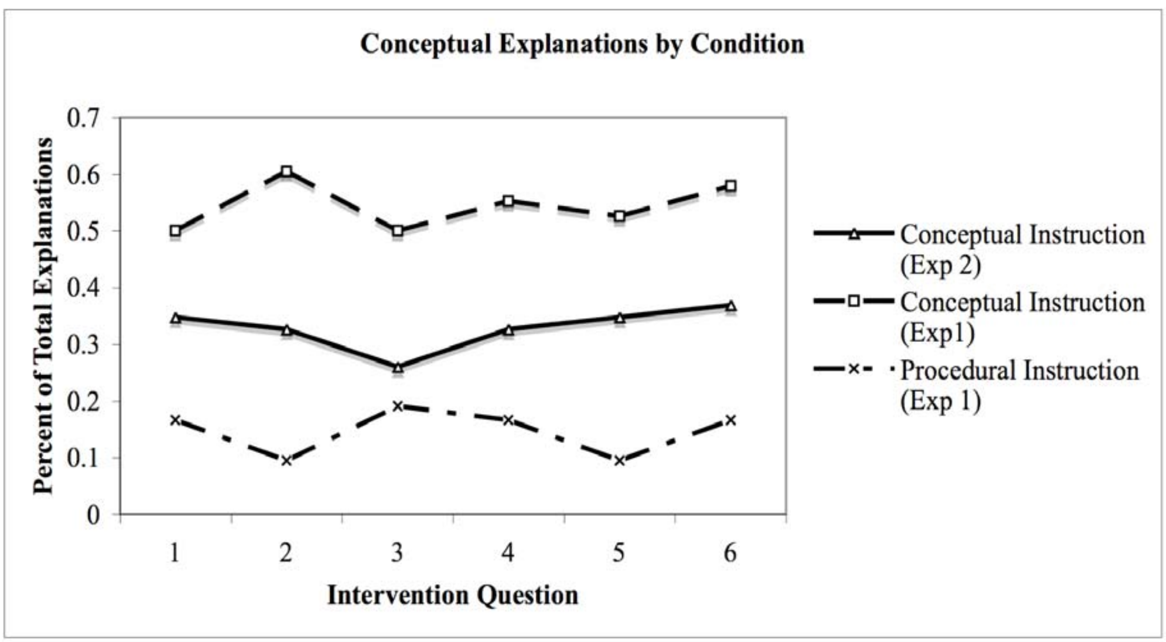


Figure 2. Proportion of Conceptual Explanations Offered by Condition

($M = .15$, $SD = .28$), $F(1,39) = 15.29$, $p < .01$, $\eta^2 = .29$. Similarly, children in the procedural instruction condition provided a procedural rationale ($M = .53$, $SD = .33$) much more frequently than those in the conceptual condition ($M = .05$, $SD = .09$), $F(1,39) = 37.81$, $p = .01$, $\eta^2 = .50$. Sixteen of eighteen students in the conceptual condition used a conceptual explanation at least once, whereas only 8 of 21 in the procedural condition did so, χ^2 (1, 39) = 10.57, p < .01. Hence, the data support our hypothesis that the type of instruction would differentially lead children to think explicitly about the conceptual rationale underlying the problems in response to prompts.

Posttest and Retention Test

We expected equivalent performance on the procedural learning problems across conditions, but greater performance on procedural transfer and conceptual items for the conceptual instruction condition. To evaluate this, we conducted a series of repeated measures ANCOVAs for procedural learning, procedural transfer, and conceptual knowledge scores, respectively, with time of assessment (posttest vs. retention) as a within-subject factor and type of instruction as a between-subject factor. Again, procedural and conceptual pretest scores and grade level were included as covariates to control for prior knowledge differences. In later analyses, we included frequency of conceptual explanations during the intervention to explore the role of explanation quality in predicting learning outcomes.

Procedural knowledge. Procedural learning was similar across conditions, $F(1, 35) = .03$, $p = .87$, $\eta^2 = .00$, and did not depend on prior knowledge (see Figure 3). There was some forgetting from posttest to retention, $F(1, 35) = 7.37$, $p = .01$, $\eta^2 = .12$, though there was no difference in forgetting across instructional conditions, $F(1, 35) = 1.20$, $p = .28$, $\eta^2 = .03$. Contrary to our expectations, however, procedural transfer was also similar across conditions, $F(1, 35) = .91$, $p = .35$, $\eta^2 = .03$, and did not depend on prior knowledge. As with learning, there was some forgetting from posttest to retention, $F(1, 35) = 9.17$, $p = .01$, $\eta^2 = .21$, but no difference in forgetting across instructional condition, , $F(1, 35) = .05$, $p = .82$, $\eta^2 = .00$. Although children in the conceptual instruction condition were never given explicit exposure to a solution procedure, they were still able to generate and transfer correct solution procedures. However, unlike past research comparing procedural and conceptual instruction – which did not include self-explanation prompts – procedural instruction was as effective as conceptual instruction at supporting procedural transfer.

Conceptual Knowledge. As expected, conceptually instructed students showed superior gains in conceptual knowledge, $F(1,35) = 16.11$, $p < .01$, $\eta^2 = .32$. This effect was over and above the main effect of prior conceptual knowledge, $F(1,35) = 6.94$, $p = .01$, $\eta^2 = .17$. None of the other variables was significant. All told, when compared to procedural instruction, conceptual instruction led to equivalent gains in procedural knowledge and superior gains in conceptual knowledge.

Explanation quality as a predictor of learning. We expected explanation quality to predict learning gains. To evaluate this portion of our hypothesis, we conducted repeated measures ANCOVAs similar to those reported above, with the exception that the frequency of conceptual explanations was included in the analysis.

The frequency of conceptual explanations was predictive of learning outcomes on all three measures: for procedural learning problems, $F(1,34) = 8.22$, $p = .01$, $\eta^2 = .20$

procedural transfer problems, $F(1,34) = 12.48$, $p < .01$, $\eta^2 = .27$, and for conceptual knowledge, $F(1,34) = 8.08$, $p = .01$, $\eta^2 = .19$. We found this positive relation after controlling for type of instruction and prior knowledge, suggesting that neither the similarity in our criteria for conceptual explanations and conceptual instruction, nor prior conceptual knowledge, accounted for this relation. Condition continued to significantly predict conceptual knowledge gains when frequency of conceptual explanations was added to the analysis, $F(1,34) = 7.40$, $p = .01$, $\eta^2 = .18$, although the portion of variance explained by condition fell from 32% to 18%. This suggests that improving explanation quality partially accounted for the effect of conceptual instruction on conceptual knowledge. More broadly, these findings are in accord with previous work showing that quality of explanation predicts learning (e.g. Renkl, 1997).

Overall, the data from Experiment 1 indicate that, when compared to procedural instruction, conceptual instruction promoted more conceptual explanations, similar gains in procedural knowledge, and superior gains in conceptual knowledge. In contrast to past research that did not include prompts for self-explanation, we found procedural instruction to be as effective as conceptual instruction at supporting procedural transfer. This may suggest that self-explanation does do something to augment the effects of procedural instruction. Performance during the intervention suggested that conceptual instruction reduced accuracy and increased variability, consistent with the hypothesis that this condition required greater search during initial problem solving. Conceptual instruction also promoted more conceptually oriented explanations. In turn, conceptually oriented self-explanations predicted gains in all three outcomes, even when controlling for type of instruction and prior knowledge. Because differences in explanation predicted

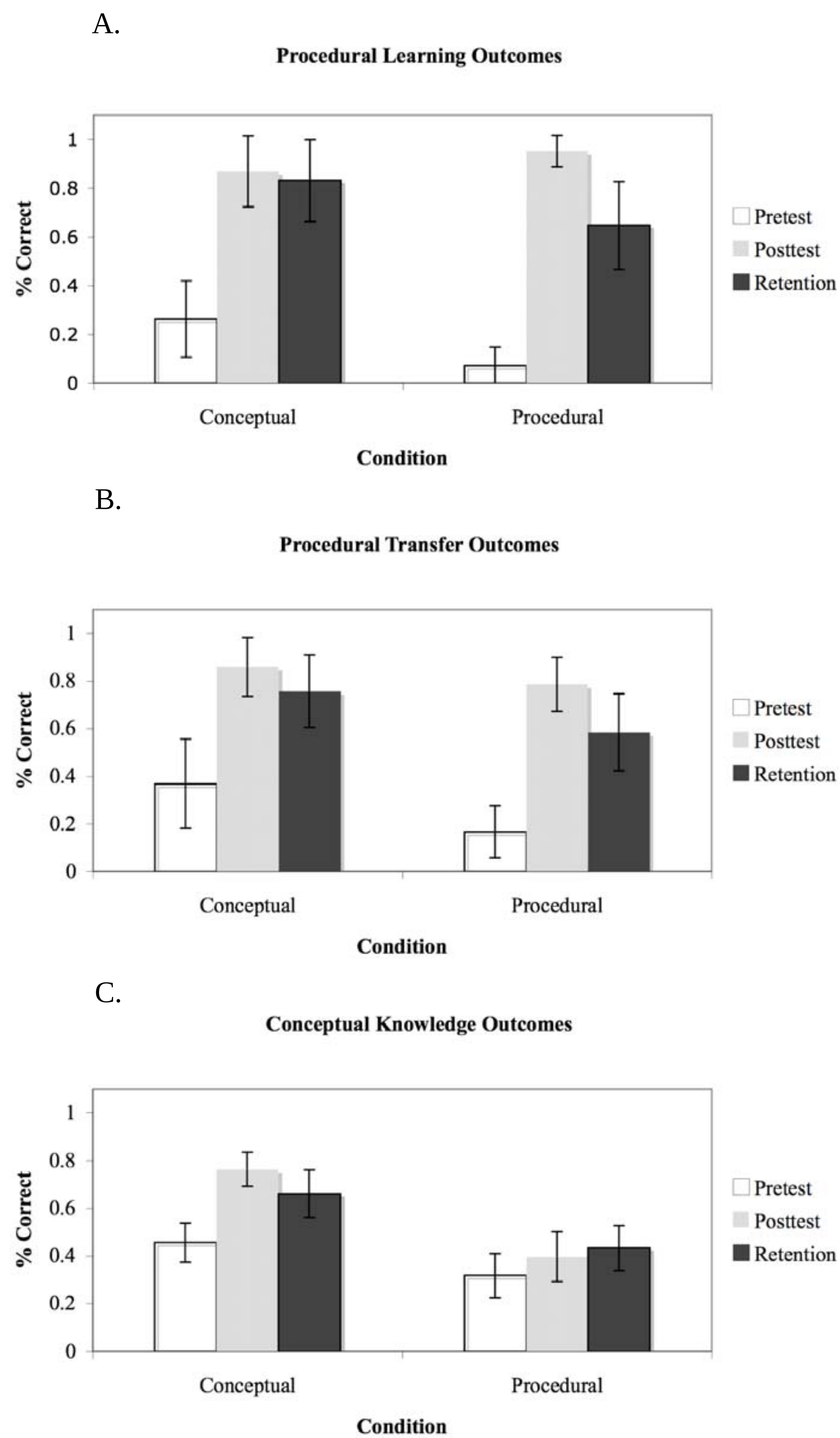


Figure 3. Accuracy on Procedural and Conceptual Knowledge Assessments Experiment 1

performance, independent of condition and prior knowledge, it seems that these differences reflect differences in the way students thought about the conceptual rationale underlying the problems.

EXPERIMENT 2

Experiment 1 demonstrated that quality of explanation varied by type of instruction and predicted gains in both conceptual and procedural knowledge over and above the effects of condition and prior knowledge. Because all students received self-explanation prompts, however, it was unclear what role the prompts played in promoting learning. It could be that conceptual instruction alone was responsible for the differences in learning gains, independent of prompts for self-explanation.

Rittle-Johnson (2006) showed that self-explanation prompts improved procedural transfer independent of procedural instruction on math equivalence problems, although they did not improve conceptual knowledge. It is unknown, however, if such effects for self-explanation prompts should be expected when used with conceptual instruction. Several studies by Chi have found effects for self-explanation prompts when used with what could reasonably be considered conceptual instruction, but the nature of the domains were significantly different from the present tasks (Chi et al, 1994; Chi, 2000; Chi & VanLehn, 1992). Chi's chosen domains are arguably much more complicated, and the method of instruction generally involves reading expository text from science books. These differences in both the tasks and the method of instruction raise questions about whether those findings generalize to our problem-solving task. We conducted Experiment 2 to evaluate the effect of self-explanation prompts in combination with conceptual instruction for problem solving. Specifically, we manipulated self-explanation prompts to see if similar performance to that of the prompted, conceptually instructed groups in Experiment I would be promoted by conceptual instruction without accompanying self-explanation prompts.

We also worked to equate time on task between conditions, an important factor that has rarely been controlled in prior research on the self-explanation effect. The vast majority of prior experimental studies have held the number of examples or problems studied constant, with the result that students in the self-explanation conditions spend more time on the intervention (e.g. Atkinson, Renkl & Merrill, 2003; Pine & Messer, 2000; Rittle-Johnson, 2001; Siegler, 1995; Wong, Lawson & Keeves, 2002). Given that generating self-explanations generally requires more time per problem, it may be that self-explanation effects arise simply from encouraging students to spend more time thinking about the material, rather than by some mechanism specific to self-explanation.

The extant evidence addressing the issue is inconclusive. Of the four studies we found that equated time on task, two found a benefit for self-explanation prompts (Aleven & Koedinger, 2002; de Bruin, Rikers & Schmidt, 2007), and two did not (Große & Renkl, 2003 Mwangi & Sweller, 1998). Moreover, these studies controlled for time in different ways. Two measured the amount of time taken by students in the self-explain condition and required that students in the no-explain control conditions spend just as long studying the same problems (de Bruin, Rikers & Schmidt, 2007; Mwangi & Sweller, 1998). Another imposed a time limit on all participants *a priori* – without regard for how long it took to complete self-explanation – which meant that some self-explanations might have been rushed or incomplete (Große & Renkl, 2003). Finally, the manipulation of Aleven & Koedinger (2002) came closest to ours, controlling for time on task by allowing participants in the no-explain condition to complete more practice problems. This study, however, differed from most self-explanation studies in that students 1) were explicitly instructed to reference a glossary containing conceptual information and 2) received feedback on their explanations. These disparate findings and methodologies cloud evaluation of whether or not self-explanation prompts are more effective than additional study time or problem solving practice. The current study seeks to help clarify this issue by equating time on task in the *explain* and *no-explain* conditions. In Experiment 2, we expected to find effects for self-explanation prompts on two of our three post-test measures when all students received conceptual instruction. We expected that conceptual instruction with practice and feedback would promote procedural learning, especially because students had additional practice opportunities. We further expected that adding self-explanation prompts would lead to even greater transfer and conceptual knowledge gains. With regard to procedural transfer, self-explanation prompts should help students make stronger links between their prior knowledge, the procedures they generate during intervention, and the concepts that govern the domain generally. Thus, self-explanation prompts should help buttress correct and flexible strategy use, improving procedural transfer performance. With regard to conceptual knowledge, we expected for self-explanation prompts to lead to continued and more productive use of the principled thought already primed by conceptual instruction.

Method

Participants

Consent was obtained from 98 third- through fifth-grade children from an urban parochial school serving a middle-class, predominantly Caucasian population. A pretest was given to identify children who could not already correctly solve half of the math equivalence problems targeted for intervention.[2] The final sample consisted of 48 children: 24 third graders (12 girls), 16 fourth graders (10 girls), and 8 fifth graders (4 girls). Their average age was 9.3 years (range 7.2-11.1). One additional child was dropped from the study for failing to complete the intervention due to emotional duress. Children participated in the fall semester.

Design and procedure

The design and procedure were identical to that of Experiment 1 with several exceptions. All children received the conceptual instruction as provided in Experiment 1. Children were then randomly assigned to either a *self-explain* (n = 23) or *no explain* condition (n = 25). Children in the *self-explain* condition solved the same six problems and received the same self-explanation prompts as those in Experiment 1. Children in the *no explain* condition were not prompted to explain and solved both the same initial six problems and an additional six problems to help equate time on task (six A+ and six +C problems total). Total time spent on the intervention problems was similar across the *self-explain*, (M = 12.54 min, SD = 2.89 min) and *no explain* conditions (M = 12.01 min,

[2] Our test-out rate was much higher for 4th and 5th graders than would be predicted by previous studies (i.e. Alibali, 1999; Perry, Church, & Goldin-Meadow, 1988; Rittle-Johnson, 2006). Conversations with teachers indicated that current student texts deal more explicitly with the concept of equality than past versions.

SD = 4.60 min, $t(46) = .47$, $p = .64$). All assessments, scoring methods, and coding schemes were identical to those of Experiment 1.

Independent raters coded 20% of participants' procedure use across all phases of the study and their *why* explanations during the intervention. Inter-rater agreement ranged from 93% for self-explanation quality to 90% for procedure use during the intervention.

Results and Discussion

As with experiment I, we first summarize the participants' knowledge base at pretest. We then compare behavior during the intervention of children in each condition, including their accuracy, procedure use and self-explanation quality. Finally, we report on the variables that affect posttest and retention performance. Effect sizes are reported as partial eta squared (η^2) values.

Pretest

Children included in the study began with little knowledge of correct procedures for solving mathematical equivalence problems at pretest. Most (81%) did not solve any of the pretest problems correctly, 10% solved only one problem correctly, and 8% solved two problems correctly. At pretest, children typically added all four numbers or added the three numbers before the equal sign. There were no differences in accuracy across the different conditions, $F(1, 46) = .49$, $p = .49$, $\eta^2 = .01$ (see Table 3). Both groups demonstrated equivalent conceptual knowledge at pretest, $F(1,47) = .06$, $p = .80$, $\eta^2 = .00$.

Intervention

We evaluated students' accuracy, procedure use, and explanation quality during the intervention. We expected students in the self-explain condition to show superior accuracy on the first six problems and explored whether additional practice on six additional problems improved accuracy in the no explain condition. We also explored whether self-explanation prompts impacted the types or variety of procedures children invented. Finally, we characterized the quality of children's self-explanations for those who were prompted.

Accuracy. Because students solved different numbers of problems by condition, our intervention analysis was broken into two components. First, we compared the accuracy of student performance on the first six problems in both conditions. There was no difference in accuracy for students in the *explain* (M = 3.04, SD = 2.72) and *no explain* conditions, (M = 2.80, SD = 2.70), $F(1, 43) = .00$, $p = .97$, $\eta^2 = .00$.

Next, we compared the mean accuracy of student performance in the *no explain* condition on the last six problems of the intervention to their performance on the first six problems. There was a significant difference between performance on the first six (M = 2.80, SD = 2.70) and last six problems for students in the *no explain* condition (M = 3.68, SD = 2.44), $F(1, 21) = .28$, $p = .61$, $\eta^2 = .01$. Thus, the additional practice seems to have helped students in the *no explain* condition.

Procedure Use. As in Experiment 1, students invented a variety of correct procedures during the intervention. The two conditions did not differ in the frequency of use of each correct procedure. Only a minority of students used multiple correct solutions. Students in the no-explain condition were somewhat more likely to use multiple correct solutions, but this difference was not statistically significant (13% vs. 28% of children in the *explain* versus *no explain* conditions, respectively, χ^2 (1, 48) = 1.63, p =.20). Interestingly, all students in the *no explain* group who used multiple correct solutions after twelve problems did so in the first six problems of the intervention.

Additional practice seems to have helped unsuccessful students in the *no explain* condition primarily by leading to their discovery of the grouping procedure (see Table 1 for a description). Students in this condition were much more likely to have used the grouping procedure at least once after having finished twelve problems (60% of students) than after the first six problems (32% of students). A paired sign test showed the difference to be significant, $p < .01$. This additional practice, however, did not increase the *no explain* students' likelihood of using multiple correct solutions, as the discovery was made primarily by students who had failed to employ any correct solution on the first six problems.

Explanation quality. Children in the *self-explain* condition were prompted to self-explain in identical fashion to those from Experiment 1. Analysis of their self-explanations reveals that they provided a conceptual explanation on about a third of all explanations (M = .33, SD = .38, see Figure 2). Further, 14 of 23 students in the *self-explain* condition used a conceptual explanation at least once. It is interesting to note that the 33% rate is relatively low compared to the rate of conceptual explanations offered by the conceptually instructed students in Experiment 1 reported above (54%).

Posttest and Retention Test

We expected equivalent performance across conditions on the procedural learning problems, but for the explain condition to lead to greater procedural transfer and

Table 3 – Coding for procedures students used to solve math equivalence problems in Experiment 2

	Explain Condition				No Explain Condition				
Procedure	Pre	Intervention	Post	Retention	Pre	Intervention Problems 1-6	Intervention Problems 7-12	Post	Retention
Correct Procedures									
Equalize	5	25	9	13	5	29	29	18	30
Add-subtract	2	6	8	7	3	2	2	4	3
Grouping	7	20	20	9	4	15	28	13	12
Ambiguous	5	7	13	19	4	3	5	12	11
Used *Any* Correct Procedure	***19***	***58***	***50***	***48***	***16***	***49***	***64***	***47***	***46***
Incorrect Procedures									
Add all	12	10	5	5	23	7	1	15	6
Add to equal sign	35	5	7	14	30	10	7	9	15
Don’t Know	4	1	6	3	7	2	0	8	4
Other	30	26	32	30	26	31	29	23	21
Used *Any* Incorrect Procedure	***81***	***42***	***50***	***52***	***86***	***50***	***37***	***55***	***46***

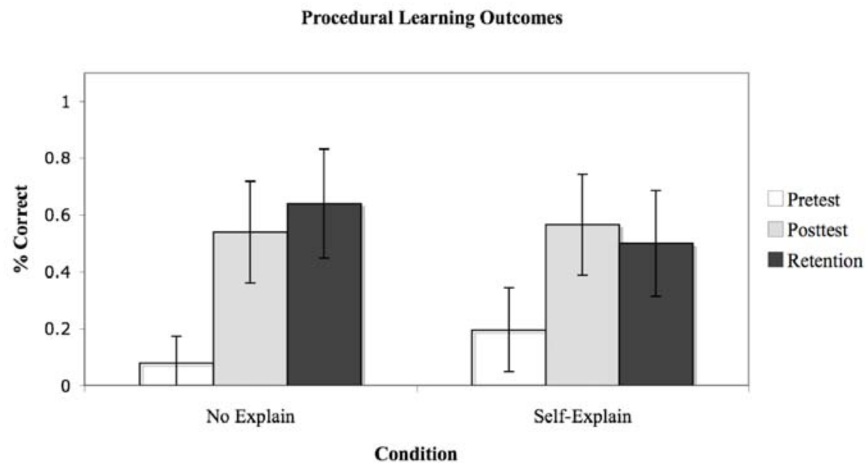
conceptual knowledge. To evaluate this, a series of repeated measures ANCOVAs was conducted for procedural learning, procedural transfer, and conceptual knowledge scores, with time of assessment (posttest vs. retention) as a within-subject factor and self-explanation prompt (present vs. absent) as a between-subject factor. Again, the prior knowledge control variables noted above were included as covariates.

Procedural knowledge. Students in both conditions demonstrated similar accuracy on procedural learning items, $F(1,43) = .44$, $p = .51$, $\eta^2 = .01$ (see Figure 4). Prior conceptual knowledge was the only variable found to have an effect on procedural learning, $F(1,43) = 4.25$, $p = .05$, $\eta^2 = .09$. Contrary to our expectations, students in both conditions also demonstrated similar accuracy on procedural transfer items, $F(1, 43) = .23$, $p = .63$, $\eta^2 = .01$. There was a trend toward conceptual knowledge predicting performance, $F(1, 43) = 3.88$, $p = .06$, $\eta^2 = .08$, with higher conceptual knowledge at pretest associated with higher procedural transfer. There was also a main effect for time, such that students showed a small improvement from the immediate posttest to the retention test, $F(1, 43) = 3.88$, $p < .01$, $\eta^2 = .16$. There were no other main effects or interactions.

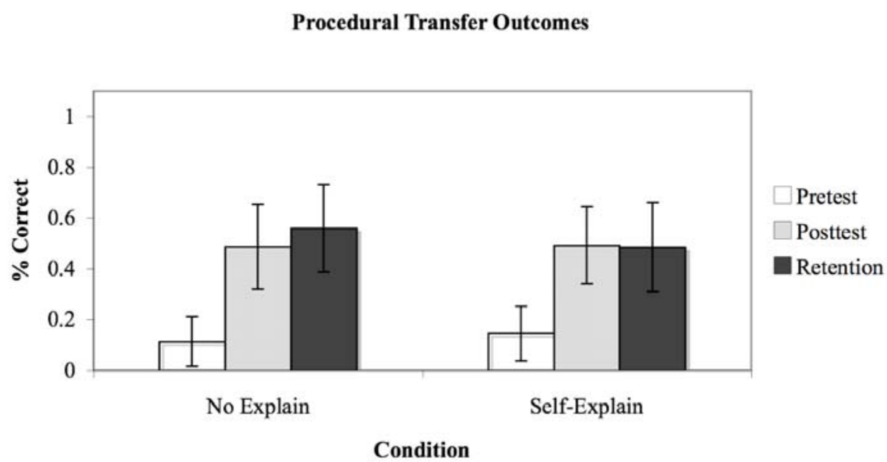
Conceptual knowledge. We also failed to find an effect for self-explanation prompts on conceptual knowledge, $F(1,43) = .63$, $p = .43$, $\eta^2 = .02$. As with our first two knowledge measures, only prior conceptual knowledge predicted knowledge gains, $F(1,43) = 11.41$, $p < .01$, $\eta^2 = .21$.

In sum, the results of Experiment 2 revealed no effects for self-explanation prompts on procedural learning, procedural transfer or conceptual knowledge when all students received conceptual instruction. Condition accounted for less than 2% of the variance in each of the outcome variables, suggesting that the null condition findings were not due to insufficient statistical power. Prior conceptual knowledge was the only factor that affected any of our three outcome measures. These data suggest that gains for students in the conceptual condition of Experiment 1 may have been due to the type of instruction, independent of self-explanation prompts.

A.



B.



C.

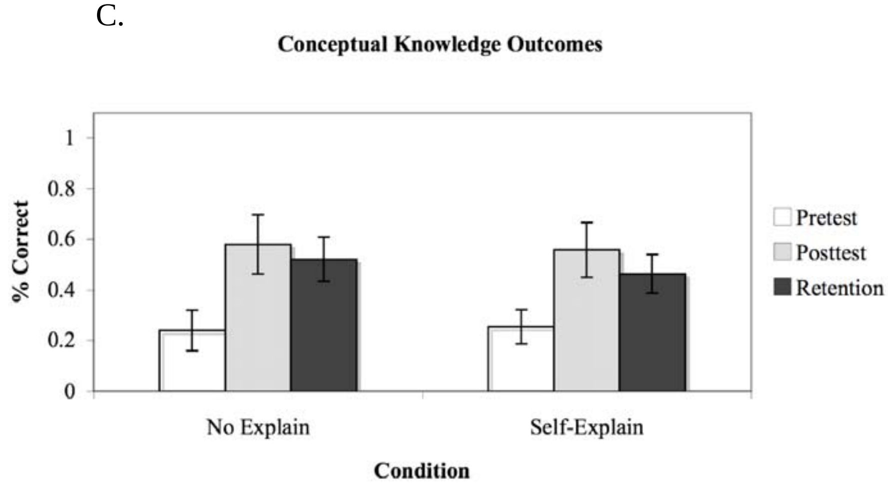


Figure 4. Accuracy on Procedural and Conceptual Knowledge Assessments Experiment 2

GENERAL DISCUSSION

In Experiment 1 we found that, compared to procedural instruction, conceptual instruction on the meaning of the equal sign promoted similar procedural knowledge and superior conceptual knowledge gains when all students self-explained. Students in the conceptual instruction group generated and transferred correct procedures even though they were never explicitly instructed on procedures. They also generated higher quality explanations, and such higher quality explanations predicted learning across instructional conditions. In Experiment 2, we found no effect for self-explanation prompts between conceptually instructed groups on measures of procedural or conceptual knowledge when students in the *no-explain* group were given additional problem solving practice to equate for time on task. This suggests that the comparative advantage of the conceptual instruction group over the procedural instruction group from Experiment 1 may have been due to type of instruction alone. Taken together, the data support two conclusions: 1) there is an asymmetry to the relation between conceptual and procedural knowledge; and 2) there may be constraints under which self-explanations can be effective, and conceptual instruction may push these constraints, attenuating the self-explanation effect in some circumstances.

Relations Between Procedural and Conceptual Knowledge

Past research suggests that there may be an asymmetric relationship between procedural and conceptual knowledge. Recall that instruction geared to boost conceptual knowledge may facilitate gains in procedural knowledge more than the opposite (e.g. Blöt et al, 2001; Kamii & Dominick, 1987; Rittle-Johnson & Alibali, 1999; Perry, 1991;

Hiebert & Wearne, 1996). The current data support this idea, at least when instruction is coupled with prompts for self-explanation. Direct instruction aimed at increasing conceptual knowledge led to gains in procedural as well as conceptual knowledge. In contrast, direct instruction aimed at procedural knowledge improved procedural knowledge but had less of an impact on conceptual knowledge. Both conceptual and procedural knowledge seem to influence each other, but the supportive effect of conceptual knowledge on procedural knowledge appears to be stronger.

Although these conclusions follow from manipulating *types of instruction*, it seems reasonable to view the asymmetry in terms of the relations between *types of knowledge*. The boost that conceptual instruction gives to conceptual knowledge seems direct – in some sense, it is teaching to the conceptual assessment. Procedural knowledge gleaned from the conceptual instruction, however, had to be generated in a secondary manner. The same argument applies to procedural instruction and the knowledge it promotes. Overall, there do seem to be asymmetrical relations between knowledge types, even in combination with self-explanation prompts.

Type of Instruction and the Mechanisms Self-Explanation

Type of instruction also may constrain the benefits of prompting for self-explanation. In Experiment 2, self-explanation prompts failed to stimulate additional knowledge gains compared to conceptual instruction alone when extra problem-solving practice was used to equate time on task. These data show that, in certain cases, conceptual instruction alone promotes conceptual and procedural knowledge acquisition as effectively as conceptual explanation coupled with prompts to self-explain. We propose three methods by which conceptual instruction may render self-explanation prompts unnecessary.

First, conceptual instruction may help children build sufficiently rich mental models that self-explanation is no longer needed. Chi et al. (1994) argue that self-explanation operates by aiding students in the repair of faulty mental models. In a parallel argument, Siegler (2002) posits that self-explanation works by getting students to consider the reasoning – particularly rule based reasoning – behind correct answers. To the extent that conceptual instruction leads to correct mental models, it may leave less room for repair, attenuating the effects of subsequent self-explanation prompts.

Second, conceptual instruction may render prompts unnecessary by encouraging more spontaneous self-explanation even in the absence of explicit prompting. Some procedural strategy must be generated to solve the problems, and it has been proposed that metacognitive processes are engaged when the existing procedural repertoire is insufficient (see Crowley, Siegler, & Shrager, 1997). These metacognitive processes may be similar to self-explanation. Hence, because conceptually instructed students are not offered a correct procedure, they may in effect engage in unprompted self-explanation to generate procedures spontaneously.

Procedural instruction, in contrast, could discourage spontaneous self-explanation. Suggestions in extant data that such instruction can diminish the attention paid to general domain principles provides some tangential evidence for this view (see, for example, evidence reported by Hiebert & Wearne, 1996; Kamii & Dominick, 1987; Perry, 1991). In particular, a robust, instructed procedure may become so successful that it obviates the need to activate the metacognitive processes posited above. In this case, self-explanations may be required to activate metacognitive processes that would lie dormant if procedural instruction were used alone. Such a mechanism would help explain previous findings of an *independent* effect for self-explanation when used in conjunction with procedural instruction (Rittle-Johnson, 2006), while perhaps rendering self-explanation prompts impotent in the current case of conceptual instruction.

There is a subtle but important distinction between the two mechanisms proposed above. On the first view, conceptual instruction may render explanation prompts ineffective because it helps students build such robust mental models that there is little repair work left for self-explanation to do. On the second view, there is work for self-explanation to do, but conceptual instruction can motivate spontaneous self-explanation without explicit prompting. The primary difference between the two alternatives lies in the amount of activity required *on the part of the learner* in constructing the final mental model. Assessment of conceptual knowledge after instruction but prior to problem solving practice would help to test roles these alternative mechanisms may play.

Finally, a third, more general, mechanism may partially account for the previously-reported benefits of self-explanation: the generation of self-explanations generally increases time spent on problem-solving tasks. Equating total time spent on the intervention task (by increasing the number of practice problems in the *no-explanation* condition of Experiment 2) resulted in no significant effect for self-explanation prompts. Given that prompting for self-explanations requires more time per problem than instruction without self-explanations, investigators should consider the tradeoffs in terms of the constraints that prompting for explanation imposes on the number of practice problems students can work through. In real world learning environments, the natural substitute for more time spent self-explaining problems is likely to be less time spent practicing additional problems. Our manipulation allowed students in the no-explanation condition to use this additional time to practice twice as many problems as those in the self-explain condition. It is noteworthy that this additional practice – in the same time span – allowed many students in Experiment 2 to discover the grouping strategy. This finding raises important questions about the overall efficiency of self-explanation prompts (see also Große & Renkl, 2003), such as whether the effect of prompting for self-explanation is due to self-explanation *per se* or if it is due to the additional time on task that such prompts encourage. Our findings contrast with those of Alevin & Koedinger (2002), who employed a parallel method for equating task, but agree with results found by others who employed different methods (Große & Renkl, 2003 Mwangi & Sweller, 1998). To resolve this issue, future studies should explicitly consider the amount of time on task afforded by alternative manipulations.

Limitations and Future Directions

An important caveat regarding the efficiency of our procedural instruction manipulation is that, although our procedural instruction taught a procedure directly, children probably were not given enough practice for the procedure to become fully automated. If this was actually the case, the problem solving procedure required conscious monitoring and rehearsal for its execution. Thus, even though our procedural instruction manipulation seems to have reduced problem-solving search, it still may not have maximally liberated cognitive resources. This might attenuate effects for procedural instruction, self-explanation prompts, or the interaction between the two. Future designs that allow for full automation of a new procedure can help allay these concerns. Measurement of automation for complex procedures presents significant, but not intractable, challenges. Conceivably, participants could be given sufficient repetitions with a particular algorithm that it becomes automated. The degree of automation might be measured by a student's speed and accuracy at the task while completing a verbal dual task.

As a further caveat, we realize that the generalizability of our conclusions is constrained by the specific scope of the current tasks. Our manipulations used rather short instructional protocols to teach relatively simple math problems in a one-on-one setting. Whether our findings generalize to more complex mathematical tasks or to classroom settings are important questions for future research.

Conclusion

In summary, we found that conceptual instruction was more efficient than procedural instruction when both were paired with prompts for self-explanation, because it supported gains in both procedural and conceptual knowledge. Additionally, we found that the benefits conferred by conceptual instruction may even preempt the benefits conferred by self-explanation prompts. The data suggest that conceptual instruction may sometimes be a more efficient means for supporting learning in mathematics than procedural instruction or self-explanation prompts.

REFERENCES

Aleven, V.A.W.M.M. & Koedinger, K.R. (2002). An effective metacognitive strategy: learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science, 26*, 147-179.

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35*, 127-145.

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, N.J.: Erlbaum.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked out steps. *Journal of Educational Psychology, 95*, 774-783.

Baroody, A. J., & Dowker, A. (2003). *The development of arithmetic concepts and skills: Constructing adaptive expertise.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Baroody, A.J., Feil, Y., & Johnson, A.R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education, 38*, 115-131.

Bednarz, N., & Janvier, B. (1988). A constructivist approach to numeration in primary school: Results of a three year intervention with the same group of children. *Educational Studies in Mathematics, 19,* 299-331.

Bielaczyc, K., Pirolli, P., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction, 13*, 221-252.

Blöt, 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,* 627-638.

de Bruin, A.B.H., Rikers, R.M.J.P., & Schmidt, H.G. (2007). The effect of self-explanation and prediction on the development of principled understanding of chess in novices. *Contemporary Educational Psychology, 32,* 188-205.

Burg, E., & Klein, A.S. (2001). Students' flexibility in solving two-digit addition and subtraction problems: Instruction effects. *Journal of Educational Psychology, 93,* 627-638.

Calin-Jageman, R. J., & Horn Ratner, H. (2005). The role of encoding in the self-explanation effect. *Cognition and Instruction, 23*, 523-543.

Chen. Z, & Klahr, D. (1999). All other things being equal: Children's acquisition of the control of variables strategy. *Child Development, 70,* 1098-1120.

Chi, M. T. H. (2000). Self-explaining: The dual processes of generating inference and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science,* Vol. 5, pp. 161-238.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How student study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

Chi, M.T.H., DeLeeuw, N., Chiu, M.H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439-477.

Chi, M. T. H., & VanLehn, K. (1991). The content of physics self-explanations. *The Journal of the Learning Sciences,* 1, 69-105.

Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatley, G., Trigatti, B., & Perlwitz, M. (1991). Assessment of a problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education, 22,* 3-29.

Conati, C., & VanLehn, K. A. (2000). Toward computer based support of meta-cognitive skills: A computational framework to coach self-explanation. International *Journal of Artificial Intelligence in Education, 11,* 398-415.

Crowley, K., Shrager, J., & Siegler, R.S. & (1997). Strategy discovery as a competitive negotiation between metacognitive and associative mechanisms. *Developmental Review, 17,* 462-489.

Didierjean, A., & Cauzinille Marmeche, E. (1997). Eliciting self-explanations improves problem solving: What processes are involved? *Current Psychology of Cognition*, 16, 325-351.

Fuson, K. C., & Briars, D. J. (1990). Using a base-ten blocks learning/teaching approach for first- and second-grade place value and multidigit addition and subtraction. *Journal for Research in Mathematics Education, 21,* 180-206. *Proceedings of the European Cognitive Science Conference 2003* (pp. 163-168). Mahwah, NJ: Erlbaum.

Graham, J. W., Hofer, S. M., & MacKinnon, D. P. (1996). Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research, 31*(2), 197-218.

Greeno, J. G., Riley, M. S., & Gelman, R. (1984). Conceptual competence and children's counting. *Cognitive Psychology, 16,* 94-143.

Große, C. S., & Renkl, A. (2003). Example-based learning with multiple solution methods facilitates understanding. In F. Schmalhofer, R. M. Young, & G. Katz (Eds.), Proceedings of the European cognitive science conference 2003 (pp. 163 – 168). Mahwah, NJ: Erlbaum.

Harrell, F. E. (2001). *Regression Modeling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer.

Hiebert, J., & Grouws, D. A. (2007). *Effective teaching for the development of skill and conceptual understanding of number: What is most effective?* Research Brief from NCTM. Reston, VA: National Council of Teachers of Mathematics.

Hiebert, J., & LeFevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1-27). Hillsdale, NJ: Erlbaum.

Hiebert, J., & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. *Cognition and Instruction, 14,* 251-283.

Kamii, C., & Dominick, A. (1997). To teach or not to teach algorithms. *Journal of Mathematical Behavior, 16*, 51-61.

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: evidence from children's metalinguistic and repair data. *Cognition, 23,* 95-147.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12,* 317-326.

Knuth, E.J., Stephens, A.C., McNeil, N.M., & Alibali, M.W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education, 37*, 4, 297-312.

Kotovsky, K., Hayes, J.R., & Simon, H.A. (1985). Why are some problems hard? Evidence from tower of Hanoi. *Cognitive Psychology, 17,* 248-294.

McNeil, N.M. & Alibali, M.W. (2005). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition and Development, 6,* 285-306.

Mwangi, W., & Sweller, J. (1998). Learning to solve compare word problems: The effect of example format and generating self-explanations. *Cognition and Instruction,* 16, 173-199.

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development*, 6, 449-468.

Perry, M., Church, R. B., & Goldin Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. Cognitive Development, 3, 359-400.

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research, 74*, 525-556.

Pine, K. J., & Messer, D. J. (2000). The effect of explaining another's actions on children's implicit theories of balance. *Cognition and Instruction*, 18, 35-52.

Pirolli, P., & Recker, M. (1994). Learning strategies and transfer in the domain of programming. *Cognition and Instruction*, *12*, 235-275.

Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and human performance*. Thousand Oaks, CA: Sage.

Renkl, A. (1997). Learning from worked-out examples: A study on individual Differences. *Cognitive Science, 21*, 1-29.

Renkl, A. (1999). Learning mathematics from worked-out examples: analysing and fostering self-explanations. *European Journal of Psychology of Education, 14*, 477-488.

Renkl, A. (2002). Worked-out examples: instructional explanations support learning by self-explanations. *Learning and Instruction, 12*, 529–556

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23*, 90-108.

Rittle-Johnson, B. (2006) Promoting transfer: effects of self-explanation and direct instruction. *Child Development*, *77*(1), 1-29.

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, *91*, 175-189.

Rittle-Johnson, B., & Russo, S. R. (1999). Learning from explaining: Does it matter if someone is listening? Paper presented at the biennial meeting of the Society for Research in Child Development, Albuquerque, NM.

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

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147-177.

Schworm, S. & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology, 99*, 285–296.

Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology*, 28, 225-273.

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

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

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257-285.

Van Merrienboer, J. J. G., Jelsma, O., and Paas, F. G. W. C. (1992). Training for reflective expertise: A four-component instructional design model for training complex cognitive skills. *Educational Technology Research and Development.* 40, 23-43.

Wong, R. M. F., Lawson, M. J., & Keeves, J. (2002). The effects of self-explanation training on students' problem solving in high-school mathematics. *Learning and Instruction, 12*, 233-262.