Journal of Experimental Child Psychology xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

In pursuit of knowledge: Comparing self-explanations, concepts, and procedures as pedagogical tools

Percival Matthews *, Bethany Rittle-Johnson

Department of Psychology and Human Development, Vanderbilt University, Peabody 422, 230 Appleton Place, Nashville, TN 37203, USA

ARTICLE INFO

Article history:
Received 22 May 2008
Revised 27 August 2008
Available online xxxx

Keywords:
Self-explanation
Problem solving
Conceptual knowledge
Procedural knowledge
Transfer

ABSTRACT

Explaining new ideas to oneself can promote learning and transfer, but questions remain about how to maximize the pedagogical value of self-explanations. This study investigated how type of instruction affected self-explanation quality and subsequent learning outcomes for second- through fifth-grade children learning to solve mathematical equivalence problems (e.g., 7 + 3 + 9 = 7 + _). Experiment 1 varied whether instruction was conceptual or procedural in nature ($n = 40$), and Experiment 2 varied whether children were prompted to self-explain after conceptual instruction ($n = 48$). Conceptual instruction led to higher quality explanations, greater conceptual knowledge, and similar procedural knowledge compared with procedural instruction. No effect was found for self-explanation prompts. Conceptual instruction can be more efficient than procedural instruction and may make self-explanation prompts unnecessary.

Introduction

Learning is often plagued both by a lack of connected understanding and by the inability to transfer knowledge to novel problems. Understanding the processes that affect knowledge change is central both to theories of learning and to the development of effective pedagogies for overcoming these problems. Prompting students to generate explicit explanations of the material they study has emerged as one potentially effective tool for promoting learning and transfer in numerous domains (e.g., Chi, DeLeeuw, Chiu, & LaVancher, 1994). Although prompting for such "self-explanations" has been shown to facilitate learning, little is known about how these prompts interact with different

* Corresponding author. Fax: +1 615 343 9494.
E-mail address: percival.g.matthews@vanderbilt.edu (P. Matthews).

0022-0965/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jecp.2008.08.004

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 affected the quality of self-explanations generated, and (c) whether self-explanation prompts were effective over and above conceptual instruction alone.

The self-explanation effect

In their seminal study on the self-explanation effect, Chi, Bassok, Lewis, Reimann, and 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, 2002), interest calculations (Renkl, Stark, Gruber, & Mandl, 1998), argumentation (Schworm & Renkl, 2007), Piagetian number conservation (Siegler, 1995), biology text comprehension (Chi et al., 1994), and balancing beam problems (Pine & Messer, 2000). Moreover, these self-explanation effects have been demonstrated across a wide range of age cohorts, from 4-year-olds (Rittle-Johnson, Saylor, & Swygert, 2007) 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., Renkl et al., 1998; Rittle-Johnson, 2006; Siegler, 2002).

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, to consider the goals of operators and procedures more frequently, and to show illusions of understanding less frequently (for an effective summary, see Renkl, 2002). 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. Hence, self-explanation prompts are not equally successful across learners 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 over and above other instructional scaffolds (Conati & VanLehn, 2000; Didierjean & Cauzinille Marmeche, 1997; Grobe & Renkl, 2003; Mwangi & Sweller, 1998).

Thus, although the relation between self-explanation prompts and improved learning has been documented and replicated, not all learners generate effective self-explanation even when prompted. 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 20th century (for an overview, see Baroody & Dowker, 2003), yet the relations between the types of instruction employed and the types of mathematical understandings generated remain largely unresolved. 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 current 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?

First, we offer some functional definitions. 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; Rittle-Johnson, Siegler, & Alibali, 2001: Star, 2005). Similarly, 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. Although mathematics education researchers sometimes emphasize conceptual knowledge at the expense of procedural knowledge, we recognize that both procedural and conceptual knowledge are critically important (National Mathematics Advisory Panel., 2008; Star, 2005). Greater conceptual and procedural knowledge both are associated with better performance on a variety of problem types (e.g., Blöt, Van der Burg, & Klein, 2001; Byrnes, 1992; LeFevre, Greenham, & Waheed, 1993; Rittle-Johnson et al., 2001), and both are key characteristics of expertise (Koedinger & Anderson, 1990; Peskin, 1998; Schoenfeld & Herrmann, 1982). Hence, we are ultimately interested in instruction that can maximally promote both types of knowledge.

Several classroom researchers have argued that, compared with procedural instruction, conceptual instruction supports more general knowledge gains. Hiebert and Wearne's (1996) study of place value and multidigit arithmetic is one widely cited case. Procedural instruction that focused on standard algorithms could quickly move students' procedural knowledge ahead of their conceptual knowledge. In contrast, conceptual instruction that focused on the base-10 system and inventing procedures improved both procedural and conceptual knowledge simultaneously. Others also have found evidence that, compared with procedural instruction, an emphasis on conceptual instruction leads to greater conceptual knowledge and to comparable procedural knowledge (Bednarz & Janvier, 1988; Blöt et al., 2001; Cobb et al., 1991; Fuson & Briars, 1990; Hiebert & Grouws, 2007; Kamii & Dominick, 1997).

Randomized experimental studies have provided some corroboration of these classroom findings. All of these studies have used math equivalence problems as the target task. In one early precursor to the current experiment, providing students with conceptual instruction led many children to generate accurate solution procedures that they could appropriately adapt to solve transfer problems (Perry, 1991). In contrast, procedural instruction improved performance on problems specifically targeted by instruction but was less effective in promoting procedural transfer. Similarly, Rittle-Johnson and Alibali (1999) found 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 when students who received hybrid instruction on both concepts and procedures performed worse on procedural transfer items than did those who received instruction on concepts alone.

These findings notwithstanding, we should be careful not to conclude prematurely that conceptual instruction is typically more effective than procedural instruction in promoting conceptual knowledge and procedural transfer. First, in the classroom studies considered above, some procedural instruction was typically included in the conceptual instruction and children were not randomly assigned to condition, making it difficult to draw conclusions about the effects of one type of instruction versus the other. Second, the experimental studies considered above offered few examples, little opportunity for practice or feedback, and/or no prompts for reflection, all of which may be important for establishing effects of procedural instruction (e.g., Peled & Segalis, 2005; Siegler, 2002). Third, recent experiments have shown explicitly 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).

All told, prior experimental studies often have not provided opportunities for practicing and reflecting on procedures that should reduce cognitive load, increase problem-solving efficiency, and free cognitive resources for improving procedural transfer and conceptual knowledge (e.g., Kotovsky, Hayes, & Simon, 1985; Proctor & Dutta, 1995; Sweller, 1988). The procedural instruction intervention in the current study offered both instruction on a procedure and multiple opportunities for practice using that procedure with feedback. Moreover, this study incorporates self-explanation prompts for reflection that may further boost the effects of procedural instruction. Such boosts to the efficacy of procedural instruction may make it equal to or more effective than conceptual 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.

Self-explanations can promote transfer when used in combination with procedural instruction to teach mathematical equivalence (Rittle-Johnson, 2006). Self-explanation prompts may push learners to consider the conceptual underpinnings of instructed procedures. Similarly, procedural instruction may free cognitive resources that can be dedicated to generating more effective self-explanations than when conceptual instruction is provided. Alternatively, conceptual instruction may boost the benefits of self-explanation prompts by directly augmenting knowledge of domain principles and directing attention to conceptual structure. Hence, self-explanation prompts might help students to fill in knowledge gaps by promoting inferences that can be drawn from knowledge provided by conceptual instruction. To date, direct comparison of self-explanation effects across the two types of instruction remains unexamined.

The current experiments

The current experiments investigated the relations among 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, 1991; Rittle-Johnson, 2006). Importantly, these problems tap children's understanding of equality, which is a fundamental concept in arithmetic and algebra (Kieran, 1981; Knuth, Stephens, McNeil, & Alibali, 2006; McNeil & Alibali, 2005). Because equality is such a central concept in mathematics, the current task is potentially fruitful for exploring 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).

The goals of the current 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 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 equating 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 and spent approximately the same amount of time on the intervention.

Experiment 1

We hypothesized that, during the intervention, (a) conceptual and procedural instruction would lead to different patterns of explanation quality, accuracy, and procedure use, (b) both types of instruction would lead to comparable procedural learning by posttest, and (c) conceptual instruction would promote conceptual knowledge and procedural transfer superior to those of procedural instruction, at least in part due to promoting more conceptual self-explanations.

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 solve a majority of math equivalence problems correctly. Students who solved more than half of the pretest problems correctly were excluded from the study, and the remaining students were randomly assigned to instructional condition. The final sample consisted of 40 children: 14 second graders (9 girls and 5 boys), 8 third graders (4 girls and 4 boys), 5 fourth graders (3 girls and 2 boys), and 13 fifth graders (6 girls and 7 boys). Their average age was 9.6 years (range = 7.5–11.8). Teachers indicated that most students had previously encountered math equivalence problems but that exposure was not frequent. Children participated during the spring semester.

Design

Children completed a pretest, an intervention, an immediate posttest, and a 2-week retention test. Students who solved no more than half of the pretest problems correctly were randomly assigned to either the procedural instruction condition ($n = 21$) 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.

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 + _ and 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 + _ and 5 + 7 + 3 = _ + 9), (b) had the blank on the left side of the equation (i.e., _ + 9 = 8 + 5 + 9) and 8 + _ = 8 + 6 + 4), or (c) included subtraction (i.e., 8 + 5 – 3 = 8 + _ and 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. In contrast, the transfer problem formats remained unfamiliar to the children at posttest and so had to be solved by applying or adapting procedures learned during the intervention—a standard approach for measuring transfer. Children were encouraged to show their calculations when solving the problems.

Q2 The five items on the conceptual knowledge assessment are described 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 and Alibali (1999) and were designed to measure both explicit and implicit conceptual knowledge.

Procedure

Children completed the written pretest during a 30-min session in their classrooms. Within 1 week of the pretest, each participant completed a one-on-one intervention and immediate posttest during one session lasting approximately 45 min. The intervention session was conducted by the first author in a quiet room at the school. The retention test was administered approximately 2 weeks later in a group session lasting no longer than 30 min.

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 (per Perry, 1991; Rittle-Johnson, 2006) 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,

> There's more than one way to solve this type of problem, but I'm going to show you one way to solve them today. 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 [using a marker to draw a circle around the 3 + 4 + 2] and then subtract the 3 that's over here [underlining the 3], and that amount goes in the blank. So, for this problem, what is 3 + 4 + 2? [waiting for student response] Right, 9, and 9 minus 3 is what? [waiting for student response] Great, so our answer is 6.

Please cite this article in press as: Matthews, P. & Rittle-Johnson, B. In pursuit of knowledge: Comparing self-explanations, ... *Journal of Experimental Child Psychology* (2008), doi:10.1016/j.jecp.2008.08.004

Table 1
Conceptual knowledge assessment items

Concept	Item	Scoring criteria
Meaning of equal sign	1. Define what the equal sign means	1 point if defined relationally (e.g., “equivalent to,” “same on both sides,” “the numbers on each side are balanced”)
	2. Rate definitions of equal sign: Rate each of the following four definitions as *always true, sometimes true*, or *never true:* (a) “The equal sign means count higher” (b) “The equal sign means two amounts are the same” (c) “The equal sign means what the answer is” (d) “The equal sign means the total”	1 point if student rated the statement “The equal sign means two amounts are the same” as *always true*
Structure of equation	3. Correct encoding: Reproduce four equivalence problems, one at a time, from memory after a 5-s delay[a]	1 point if student puts numerals, operators, equal sign, and blank in correct respective positions for all four problems
	4. Recognize correct use of equal sign in multiple contexts: Indicate whether eight equations (e.g., 8 = 2 + 6 and 3 + 2 = 7 − 2) make sense	1 point if more than 75% correct
Meaning of equal sign and structure of equation	5a. Record the two separate sides of the equation 4 + 3 = 5 + 2	(a) 1 point if 4 + 3 and 5 + 2 identified as separate sides of the equation
	5b. State the meaning of the equal sign in *this* problem	(b) 1 point if defined relationally as above

[a] This task is based on Larkin’s (1989) suggestion that increased conceptual knowledge results in an improved ability to “see” a hierarchical and organized structure where one exists, which in this case is the placement of the equal sign separating the equation into two sides. Past work has shown that many students reconstruct the equations with the equal sign at the end of the sentence (e.g., 4 + 5 + 7 = 4 + _ as 4 + 5 + 7 + 4 = _) and that how children reconstruct equations is related to their knowledge of equivalence (McNeil & Alibali, 2004).

After receiving instruction on the two standard A+ problems, students received similar instruction on two +C questions problems and a final A+ problem. In past instructional studies on mathematical equivalence, students have received instruction on only two instances of a single problem type (Perry, 1991; Rittle-Johnson, 2006; Rittle-Johnson & Alibali, 1999); we expected that instruction on two problem types and with a greater number of problems would increase the generalizability of the procedure. We chose to provide instruction on the add–subtract procedure both because it is a procedure that many students invent on their own (Rittle-Johnson, 2006) and because it can be described without reference to concepts, allowing us to keep the two types of instruction distinct.

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 [making a sweeping gesture under the left side] and one on the right side of the equal sign [making a sweeping gesture under the right side]. The first side is 3 + 4 [making a sweeping gesture]. The second side is 3 + 4 [making a 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 his hand back and forth].

Students were shown four other number sentences of various sorts (i.e., 4 + 4 = 3 + 5, 3 + 4 = _, 2 + 3 ○ 3 + 6[1], and 5 + 4 + 3 = 5 + _) and reminded of what the equal sign meant in each case. This brought the total number of examples to five so as to parallel the number of problems encountered in the procedural instruction condition. No solution procedures were ever discussed. As in the procedural instruction condition, the experimenter often prompted students with questions to ensure that they were attending to and understanding the instruction. Instruction took approximately 6 min in either condition.

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

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) and was originally adapted from Siegler (2002). Children saw a screen with the answers that two children at another school had purportedly given; one of these answers was correct and one was incorrect, as shown in Fig. 1. The experimenter then asked participants both how the other children got their answers and why each answer was correct or incorrect. Both questions were asked so as to highlight for children the distinction between how a procedure is employed and why it is correct or incorrect.

Children were asked to 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 were asked to explain both "how" and "why" because a past study found that when students are simply asked "why," they typically provided only descriptions of how the student solved the problem without reference to the underlying logic or concepts (Siegler, 2002). Separate "how" and "why" prompts were meant to encourage students to go beyond descriptions of procedures. We focus on answers to "why" questions as self-explanations for the analyses because preliminary analyses indicated that these answers focused more on the logic behind the mathematical manipulations than did responses to the "how" questions. Such inference about the logic of a domain is seen as integral to the functioning of self-explanation (Siegler, 2002). For the most part, in the "how" responses, students accurately described the intended incorrect procedure that was used to arrive at the incorrect answers and described their own correct procedure as the one the hypothetical child used to arrive at the correct answer. Thus, the "how" responses were not very informative.

The intervention was audiotaped and videotaped. Total time spent on the practice problems was similar across the conceptual (M = 15.72 min, SD = 3.89) and procedural (M = 14.86 min, SD = 3.29) instruction conditions, $t(38) = 0.76$, $p = .45$. Immediately following the intervention, children completed a paper-and-pencil posttest administered individually by the experimenter in the same room. Approximately 2 weeks later, students completed a delayed retention test as groups in their classrooms.

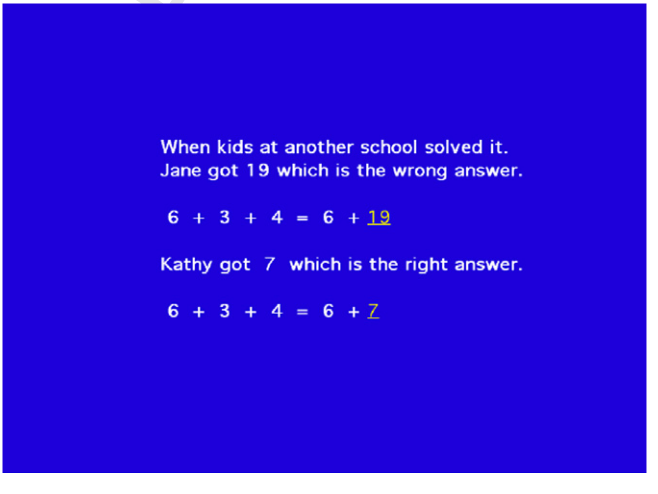

Fig. 1. Screen shot of the ~~addition screen for~~ participants in the self-explanation conditions.

Coding

Assessment. 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. On the assessments, procedure use could be inferred from children's solution and written calculations; for example, giving the answer 21 indicated use of the add all procedure to solve 7 + 3 + 4 = 7 + _, and writing 7 + 3 + 4 – 7 = 7 indicated use of the add–subtract procedure. During the intervention, children's self-reports of how they solved each problem were used to identify procedure use (see Table 2 for sample explanations. Accuracy scores were calculated based on the percentage of problems children solved using a correct procedure, regardless of whether they made arithmetic errors, so as to be consistent with past research using this task (e.g., McNeil & Alibali, 2000). Correct procedure use was coupled with arithmetic errors on only 3% of all assessment items. For conceptual knowledge items, each item was 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.

Explanation quality. Students' self-explanations of why solutions were correct and incorrect during the intervention were also 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 nonresponses (e.g., "That's what the problem tells you to do").

Independent raters coded 20% of participants' procedure use across all phases of the study and their "why" explanations during the intervention. Interrater 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 (two in the procedural instruction condition and one in the conceptual instruction condition). These partic-

Table 2
Experiment 1 procedure use: Percentage of problems on which each procedure was used, by condition

Procedure	Sample explanation	Conceptual condition				Procedural condition			
		Pre	Intervention	Post	Retention	Pre	Intervention	Post	Retention
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	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 8s 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	
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 =	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	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	

ipants did not differ from the other participants on the pretest measures. To deal with these missing data, an imputation technique was used to approximate the missing accuracy scores on the retention test. Imputation leads to more precise and unbiased conclusions than does casewise deletion (Peugh & Enders, 2004; Schafer & Graham, 2002), and simulation studies have found that using maximum likelihood (ML) imputation when data are missing at random leads to the same conclusions as when there are no missing data (Graham, Hofer, & MacKinnon, 1996; Schafer & Graham, 2002).

Because 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 [missing completely at random] test: $\chi^2(31) = 14.14$, $p > .99$). As recommended by Schafer and Graham (2002), we used the expectation maximization (EM) algorithm for ML estimation via the missing value analysis module of SPSS. Students' missing scores were estimated from all nonmissing continuous 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 casewise deletion approach (i.e., the change in η^2 was < .02 for all significant variables). There were no substantive differences between analyses conducted with casewise deletion and those conducted with imputation.

Results

First, we summarize 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 affected 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, and children typically added all four numbers or added the three numbers before the equal sign (see Table 2). There were no differences in accuracy across 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 instruction condition ($M = 46\%$, $SD = 18$) scoring somewhat 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 analyses of covariance (ANCOVAs) were conducted with type of instruction as a between-participant 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 students' 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, with children with higher procedural knowledge pretest scores being 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 to solve 97% of intervention problems (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 1 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 instruction condition were far more likely to use the add–subtract procedure than those in the conceptual instruction condition, $F(1, 35) = 36.64$, $p < .01$, $\eta^2 = .51$.

Children in the conceptual instruction condition, in contrast, employed a variety of strategies. As shown in Table 2, they used the equalize procedure most often but also used the add–subtract and grouping procedures fairly often. Not surprisingly, they were more than four times as likely to use multiple correct procedures (42 and 10% of children in the conceptual and procedural instruction conditions, respectively), $\chi^2(1, 40) = 5.65$, $p = .02$. Although they used more correct procedures, students in the conceptual instruction condition showed a strong preference for the equalizer strategy and were far more likely than those in the procedural instruction 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, the procedurally instructed group was less likely to adopt multiple correct procedures.

Explanation quality. There was a stark contrast in the explanations offered in response to the "why" questions by condition. Children who were given conceptual instruction provided a conceptual rational on more than half of all explanations ($M = .54$, $SD = .34$) (see Fig. 2), whereas children in the procedural instruction condition rarely did so ($M = .15$, $SD = .28$), $F(1, 39) = 15.29$, $p < .01$, $\eta^2 = .29$. Similarly, children in the procedural instruction condition provided a procedural rationale ($M = .53$, $SD = .33$) much more frequently than those in the conceptual instruction condition ($M = .05$, $SD = .09$), $F(1, 39) = 37.81$, $p < .01$, $\eta^2 = .50$. Also, 16 of 18 students in the conceptual instruction condition used a conceptual explanation at least once, whereas only 8 of 21 students in the procedural instruction condition did so, $\chi^2(1, 39) = 10.57$, $p < .01$. Overall, the data support our hypothesis that the type of instruction would lead to different learning pathways during the intervention, as indexed by accuracy, procedure use, and explanation quality.

Posttest and retention test

We expected equivalent performance on the procedural learning problems across conditions but greater performance on procedural transfer and conceptual knowledge 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-participant factor and type of instruction as a between-participant factor. We expected a main effect for time, with students forgetting some from posttest to retention test, but we expected the effect of condition to remain constant across posttest and retention test (i.e., no interaction between time and condition). 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, $F(1, 35) = 0.03$, $p = .87$, $\eta^2 = .00$ (see Fig. 3). Procedural learning was not predicted by either pretest procedural knowledge, $F(1, 35) = 2.34$, $p = .14$, $\eta^2 = .06$, or pretest conceptual knowledge, $F(1, 35) = 0.01$, $p = .91$, $\eta^2 = .00$. There was some forgetting from posttest to retention, $F(1, 35) = 10.59$, $p < .01$, $\eta^2 = .23$, although there was no difference in forgetting across instructional conditions, $F(1, 35) = 1.20$, $p = .28$, $\eta^2 = .03$.[2] Contrary to our expectations, procedural transfer was also similar across conditions, $F(1, 35) = 0.91$, $p = .35$, $\eta^2 = .03$, and did not depend on either pretest procedural knowledge, $F(1, 35) = 2.34$, $p = .14$, $\eta^2 = .06$, or conceptual knowledge, $F(1, 35) = 0.18$, $p = .67$, $\eta^2 = .01$. As with learning, there was some forgetting from posttest to retention, $F(1, 35) = 14.26$, $p < .01$, $\eta^2 = .29$, but no difference in forgetting across instructional conditions, $F(1, 35) = 0.06$, $p = .82$, $\eta^2 = .00$. Although children in the conceptual instruction condition

[2] Visual inspection of Fig. 2 suggests that the conceptual condition may outperform the procedural condition on the retention test. However, a follow-up univariate ANCOVA with procedural learning at retention as the dependent variable failed to indicate an effect for condition, $F(1, 35) = 0.45$, $p = .51$, $\eta^2 = .01$.

Please cite this article in press as: Matthews, P. & Rittle-Johnson, B. In pursuit of knowledge: Comparing self-explanations, ... *Journal of Experimental Child Psychology* (2008), doi:10.1016/j.jecp.2008.08.004

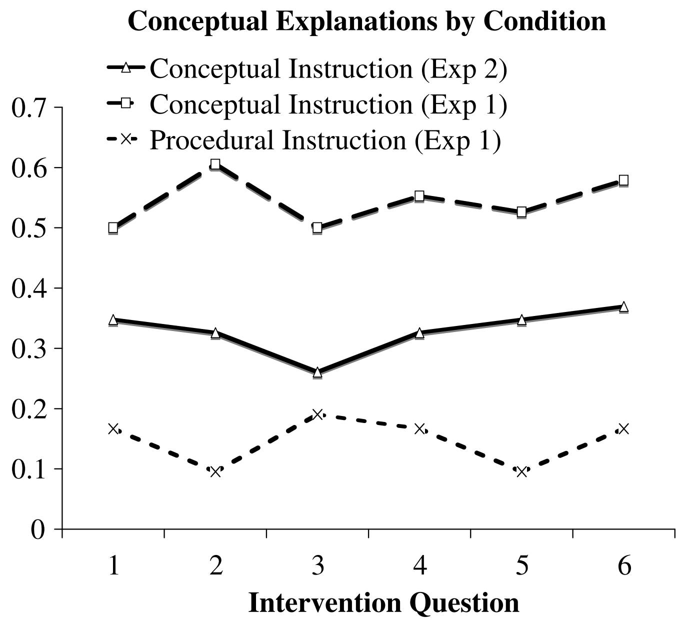

Fig. 2. Proportions of conceptual explanations offered by condition for each trial. ~~Exp, Experiment.~~

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 that did not include self-explanation prompts, procedural instruction was as effective as conceptual instruction at supporting procedural transfer. Students in the procedural instruction condition continued to use the add–subtract procedure most often, but they also used the equalize procedure 11% of the time at posttest and retention test (see Table 2). This suggests that students' strategy use at these assessment points was less rigid than during the intervention. Procedure use for the conceptual instruction condition at posttest and retention test largely mirrored that during the intervention.

Conceptual knowledge. As expected, conceptually instructed students showed superior conceptual knowledge, $F(1, 35) = 16.44$, $p < .01$, $\eta^2 = .32$. This effect was over and above the main effect of prior conceptual knowledge, $F(1, 35) = 6.96$, $p = .01$, $\eta^2 = .17$. None of the other variables was significant. All told, when compared with procedural instruction, conceptual instruction led to equivalent procedural knowledge and superior conceptual knowledge.

Explanation quality as a predictor of learning. We expected explanation quality to predict knowledge at posttest and retention test. To evaluate this, we conducted repeated-measures ANCOVAs similar to those reported above with the exception that either the frequency of conceptual explanations or the frequency of procedural explanations was included in each analysis as an additional covariate. Frequencies of the different types of explanations were analyzed separately because they were relatively highly correlated ($r = -.68$).

The frequency of conceptual explanations was predictive of learning outcomes for all three measures: procedural learning, $F(1, 34) = 8.12$, $p < .01$, $\eta^2 = .19$, procedural transfer, $F(1, 34) = 12.31$, $p < .01$, $\eta^2 = .27$, and conceptual knowledge, $F(1, 34) = 8.59$, $p < .01$, $\eta^2 = .20$. 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 when frequency of conceptual explanations, $F(1, 34) = 7.55$, $p = .01$, $\eta^2 = .18$, was added to the analysis, 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 con-

Please cite this article in press as: Matthews, P. & Rittle-Johnson, B. In pursuit of knowledge: Comparing self-explanations, ... *Journal of Experimental Child Psychology* (2008), doi:10.1016/j.jecp.2008.08.004

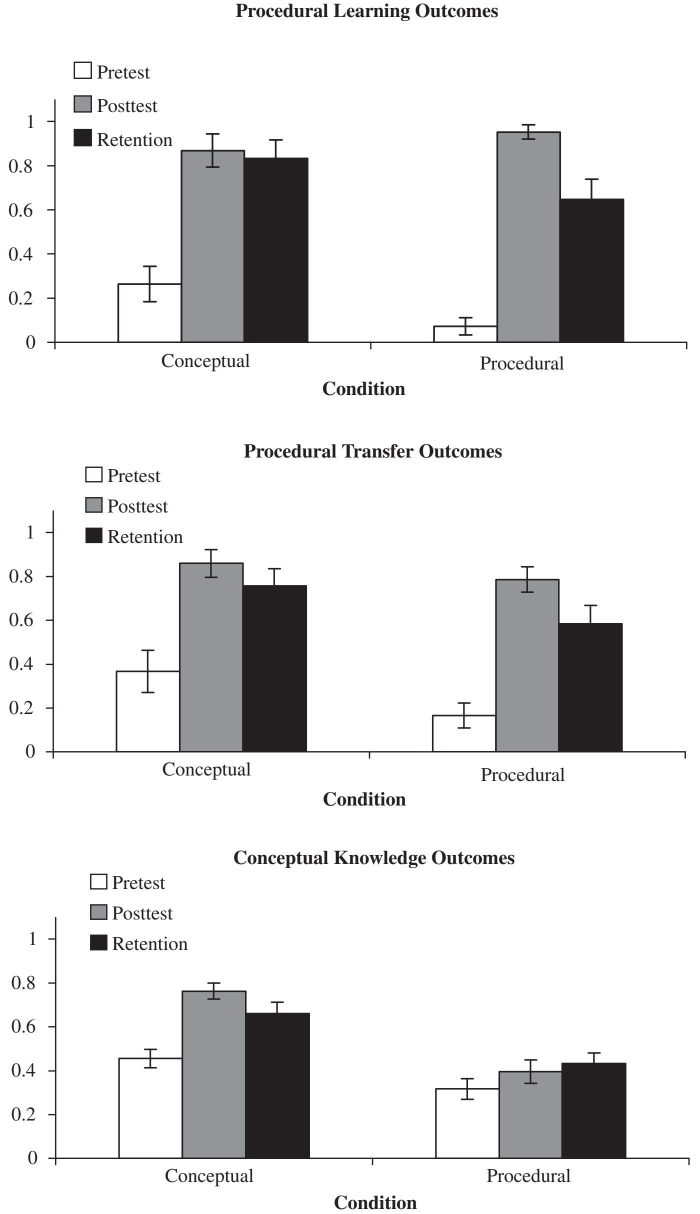

Fig. 3. Accuracy on procedural and conceptual knowledge assessments: Experiment 1. Error bars represent standard errors.

ceptual knowledge. More broadly, these findings are in accord with previous work showing that quality of explanation predicts learning (e.g., Renkl, 1997).

Frequency of procedural explanations was not predictive of procedural learning, $F(1, 34) = 0.47$ $p = .50$, $\eta^2 = .01$. It was, however, negatively predictive of conceptual knowledge, $F(1, 34) = 7.87$, $p < .01$, $\eta^2 = .19$, and showed a negatively predictive trend for procedural transfer, $F(1, 34) = 3.20$, $p = .08$, $\eta^2 = .09$. These data provide further evidence that conceptual explanations do indeed predict better overall learning than procedural explanations.

Please cite this article in press as: Matthews, P. & Rittle-Johnson, B. In pursuit of knowledge: Comparing self-explanations, ... Journal of Experimental Child Psychology (2008), doi:10.1016/j.jecp.2008.08.004

Discussion

Overall, the data from Experiment 1 indicate that, when compared with procedural instruction, conceptual instruction led to equivalent procedural knowledge and superior conceptual knowledge. Conceptual instruction also promoted more conceptual explanations. In contrast to past research that did not include prompts for self-explanation or multiple opportunities for practice, we found procedural instruction to be as effective as conceptual instruction at supporting procedural transfer. This converges with prior findings that self-explanation prompts help to improve procedural transfer when used in combination with procedural instruction (Rittle-Johnson, 2006), suggesting that self-explanation prompts supported generalization of procedures.

Performance during the intervention indicated that procedural instruction improved accuracy and reduced procedural variability, suggesting that this condition required less problem-solving search during initial problem solving. Thus, children in the procedural instruction condition had more opportunities to practice using a correct procedure, which may have also supported procedural transfer.

Although reduction in procedural variability in the procedural instruction condition allowed students to practice a correct procedure more often, use of multiple strategies is an important developmental milestone (Siegler, 1996) and is associated with greater transfer performance and greater responsiveness to instruction (Goldin-Meadow, Alibali, & Church, 1993; see also Siegler, 1996). These prior findings suggest that variability in strategy use is an important outcome, and conceptual instruction tended to support greater procedural variability. On our outcome measures, however, this increased variability in the conceptual instruction condition did not lead to greater procedural transfer. Further studies that include measures of procedural flexibility might shed more light on the impact of knowing and using multiple strategies (e.g., Star & Seifert, 2006).

We also found important differences in explanation quality across conditions. Conceptual instruction did promote more conceptually oriented explanations. In turn, conceptually oriented self-explanations were predictive of all three outcomes, even when controlling for type of instruction and prior knowledge. Because differences in explanation quality predicted 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 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 independent of self-explanation prompts.

Self-explanation prompts have been shown to improve procedural transfer independent of procedural instruction on math equivalence problems, although they did not improve conceptual knowledge (Rittle-Johnson, 2006). It is unknown, however, whether self-explanation prompts improve learning when used with conceptual instruction or under what conditions self-explanation prompts might promote conceptual knowledge for a problem-solving task. To investigate this, we manipulated the use of self-explanation prompts when all children received conceptual instruction about math equivalence. To help control for time on task, children in the no-explain condition solved twice as many practice problems as those in the self-explain condition.

In Experiment 2, we expected both groups to learn correct procedures and perform equally well on the procedural learning items. However, we hypothesized that self-explanation prompts should help students to make stronger links among their prior knowledge, the procedures they generate during the intervention, and the concepts that govern the domain generally. Thus, self-explanation prompts were expected to improve procedural transfer and conceptual knowledge even though the no-explain group studied additional problems.

Method

Participants

Consent was obtained from 98 third through fifth graders from an urban parochial school serving a middle-class, predominantly Caucasian population. A pretest was given to identify children who could not already solve half of the math equivalence problems correctly. The final sample consisted of 48 children: 24 third graders (12 girls and 12 boys), 16 fourth graders (10 girls and 6 boys), and 8 fifth graders (4 girls and 4 boys). Their average age was 9.3 years (range = 7.2–11.1). In addition, 1 child was dropped from the study for failing to complete the intervention due to emotional duress. Teachers indicated that most students had previously encountered math equivalence problems but that exposure was not frequent. Children participated during the fall semester.

Design and procedure

The design and procedure were identical to that of Experiment 1 with the following exceptions. Children who solved no more than half of the problems correctly were randomly assigned to either a *self-explain* condition ($n = 23$) or a *no-explain* condition ($n = 25$). First, all children received the conceptual instruction as provided in Experiment 1. Next, 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, $SD = 2.89$) and no-explain ($M = 12.01$ min, $SD = 4.60$) conditions, $t(46) = 0.47$, $p = .64$). 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. Interrater agreement ranged from 93% for self-explanation quality to 90% for procedure use during the intervention.

Results and discussion

As with Experiment 1, we first summarize 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.

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, and children typically added all four numbers or added the three numbers before the equal sign. There were no differences in frequency of using correct procedures across the different conditions, $F(1, 46) = 0.49$, $p = .49$, η^2 (see Table 3). Both groups also demonstrated equivalent conceptual knowledge at pretest, $F(1, 46) = 0.06$, $p = .80$, η^2. In all subsequent analyses, we controlled for conceptual and procedural knowledge scores at pretest and for grade level.

Intervention

Accuracy. Because students solved different numbers of problems by condition, our intervention analysis was broken into two components. First, we compared accuracy on the first six intervention problems between conditions. There was no difference in accuracy for students in the explain ($M = 3.04$, $SD = 2.72$) and no-explain ($M = 2.80$, $SD = 2.70$) conditions, $F(1, 43) = 0.00$, $p = .97$, $\eta^2 = .00$.

Next, we compared the mean accuracy of student performance in the no-explain condition on the last six problems of the intervention with students' 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 ($M = 3.68$, $SD = 2.44$) problems for students in the no-explain condition, $F(1, 21) = 0.28$, $p = .61$, $\eta^2 = .01$. Thus, the additional practice seems to have helped students in the no-explain condition.

YJECP 3046 ARTICLE IN PRESS No. of Pages 21, Model 1G
26 September 2008 Disk Used

Table 3
Experiment 2 procedure use: Percentage of problems on which each procedure was used, by condition

Procedure	Explain condition				No-explain condition				
	Pre	Intervention	Post	Retention	Pre	Interventions 1 to 6	Interventions 7 to 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 =	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

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 (see Table 3), and only a minority of students used multiple correct solutions (13 and 28% of children in the explain and no-explain conditions, respectively), $\chi^2(1, 48) = 1.63$, $p = .20$. Interestingly, all students in the no-explain group who used multiple correct solutions after 12 problems had done so in the first 6 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 2 for a description of the procedure). Students in this condition were much more likely to have used the grouping procedure at least once after finishing 12 problems (60% of students) than after finishing the first 6 problems (32% of students). A paired sign test showed the difference to be significant at $p < .01$. This additional practice, however, did not increase the no-explain students' likelihood of using multiple correct solutions given that the discovery was made primarily by students who had failed to employ any correct solution on the first 6 problems.

Explanation quality. Children in the self-explain condition were prompted to self-explain in identical fashion to those of Experiment 1. Analysis of their self-explanations revealed that they provided a conceptual explanation on approximately a third of all explanations ($M = .33$, $SD = .38$) (see Fig. 2). Furthermore, 14 of 23 students in the self-explain condition used a conceptual explanation at least once, and only 3 of 23 used a procedural explanation at least once.

Posttest and retention test

Procedural knowledge. As expected, students in both conditions demonstrated similar accuracy on procedural learning items, $F(1, 43) = 0.44$, $p = .51$, $\eta^2 = .01$ (see Fig. 4). Pretest conceptual knowledge was positively related to procedural learning, $F(1, 43) = 4.25$, $p = .05$, $\eta^2 = .09$, but pretest procedural knowledge was not, $F(1, 43) = 0.13$, $p = .72$, $\eta^2 = .00$. Contrary to our expectations, students in both conditions also demonstrated similar accuracy on procedural transfer items, $F(1, 43) = 0.23$, $p = .63$, $\eta^2 = .01$. There was a trend toward pretest 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. Pretest procedural knowledge did not predict procedural transfer, $F(1, 43) = 0.11$, $p = .75$, $\eta^2 = .00$. There were no other main effects or interactions. Although procedural knowledge seems to be equivalent across groups at posttest and retention test, the pattern seems to suggest that the no-explain group employed the equalize procedure more frequently (see Table 3).

Conceptual knowledge. Self-explanation prompts did not improve conceptual knowledge, $F(1, 43) = 0.63$, $p = .43$, $\eta^2 = .02$. Only pretest conceptual knowledge predicted later conceptual knowledge,

Please cite this article in press as: Matthews, P. & Rittle-Johnson, B. In pursuit of knowledge: Comparing self-explanations, ... Journal of Experimental Child Psychology (2008), doi:10.1016/j.jecp.2008.08.004

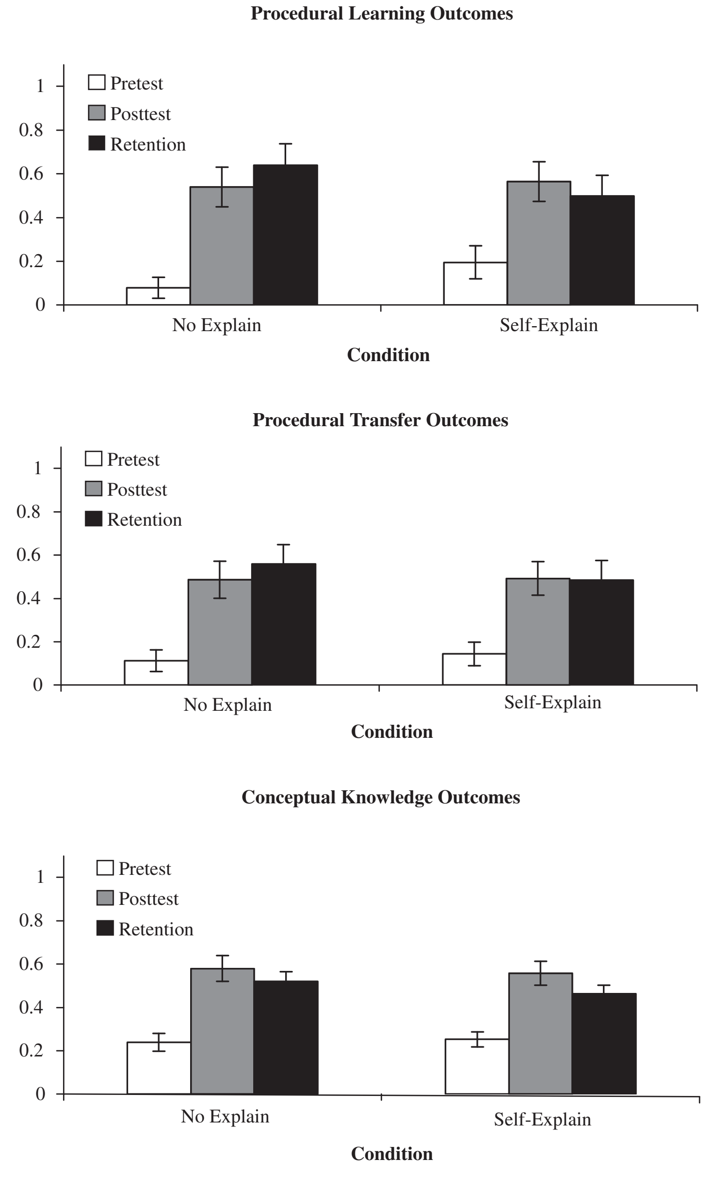

Fig. 4. Accuracy on procedural and conceptual knowledge assessments: Experiment 2. Error bars represent standard errors.

$F(1, 43) = 11.41$, $p < .01$, $\eta^2 = .21$. There was also some forgetting over time, $F(1, 43) = 5.64$, $p = .02$, $\eta^2 = .12$. There were no other main effects or interactions.

Explanation quality as a predictor of learning. We expected explanation quality to predict knowledge at posttest and retention test for the self-explain group. To evaluate this, we conducted repeated-measures ANCOVAs similar to those reported above for the students in this condition with the exception that the frequency of conceptual explanations was included in the analyses.

The frequency of conceptual explanations was predictive of procedural learning, $F(1, 18) = 6.29$, $p = .02$, $\eta^2 = .26$, and showed a trend for predicting procedural transfer, $F(1, 18) = 3.74$, $p = .07$, $\eta^2 = .17$. The frequency of conceptual explanations was not, however, predictive of conceptual knowl-

Please cite this article in press as: Matthews, P. & Rittle-Johnson, B. In pursuit of knowledge: Comparing self-explanations, ... *Journal of Experimental Child Psychology* (2008), doi:10.1016/j.jecp.2008.08.004

edge, $F(1, 18) = .06$, $p = .82$, $\eta^2 = .00$, partially because pretest conceptual knowledge accounted for most of the variance in later conceptual knowledge, $F(1, 18) = 7.74$, $p = .01$, $\eta^2 = .30$. Because all students received conceptual instruction, any boost in conceptual explanations due to type of instruction could no longer be evaluated against a comparison group.

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 3% of the variance in each of the outcome variables, suggesting that the null findings for condition were not due to insufficient statistical power. Prior conceptual knowledge was the only factor shown to affect any of our three outcome measures.

General discussion

Compared with procedural instruction, conceptual instruction on the meaning of the equal sign promoted similar procedural knowledge and superior conceptual knowledge when all students self-explained in Experiment 1. 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, which in turn predicted learning. In Experiment 2, self-explanation prompts did not improve procedural or conceptual knowledge when all students received conceptual instruction, and students in the no-explain group were given additional problem-solving practice to equate for time spent thinking about the problems. This suggests that the benefits of conceptual instruction may sometimes supplant the utility of self-explanation prompts. Taken together, the data support two conclusions. First, there is an asymmetry to the relations between conceptual and procedural knowledge. Second, 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, such that conceptual knowledge tends to support growth in both types of knowledge, whereas procedural knowledge primarily supports growth in only one type of knowledge. Recall that several investigators have found that instruction geared at boosting conceptual knowledge can facilitate gains in both conceptual and procedural knowledge, whereas instruction geared at procedural knowledge is less effective at promoting conceptual knowledge (e.g., Blöt et al., 2001; Hiebert & Wearne, 1996; Kamii & Dominick, 1997; Perry, 1991; Rittle-Johnson & Alibali, 1999). 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 both procedural and conceptual knowledge. In contrast, direct instruction aimed at procedural knowledge improved procedural knowledge but had less of an impact on conceptual knowledge. Furthermore, prior conceptual knowledge sometimes predicted later procedural knowledge, but the reverse did not occur. The supportive effect of conceptual knowledge on procedural knowledge appears to be stronger than the reverse.

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, needed 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.

Before concluding that procedural instruction is less effective than conceptual instruction, future research must investigate the effects of instruction under at least three additional conditions. First, additional problem-solving practice to allow for automation of procedures should free additional cognitive resources for reflection on the underlying concepts, and this might facilitate learning in the procedural instruction condition. For example, Baroody, Ginsburg, and Waxman (1983) found that children recognized the relation between addition and subtraction first with the well-known addition doubles (e.g., 6 + 6 = 12, so 12 − 6 = 6). Second, tasks with more complicated procedures that are difficult to invent may require procedural instruction. For instance, the invert and multiply strategy for dividing fractions is a powerful tool that is generally taught procedurally. Indeed cognitive load theory suggests that sometimes goal-free application of procedures may be of greater benefit than more conceptually based approaches (see Sweller, van Merrienboer, & Paas, 1998). Third, our study is based on one-on-one scripted instruction. These effects still need to be evaluated in ecologically valid classroom settings where interactions with teachers and peers might alter the effects.

Type of instruction and self-explanation

The results of Experiment 2 suggest that self-explanation prompts did not augment the effects of conceptual instruction on learning and transfer. In contrast, self-explanation prompts have been shown to improve procedural learning and transfer over and above procedural instruction for a comparable population using the same task (Rittle-Johnson, 2006). These divergent results suggest that the benefits of self-explanation may vary with type of instruction. In certain cases, conceptual instruction alone can be sufficient to promote conceptual and procedural knowledge acquisition as effectively as conceptual explanation coupled with prompts to self-explain, whereas procedural instruction does not seem to replace the benefits of self-explanation. We propose two pathways by which conceptual instruction may render self-explanation prompts unnecessary and then consider reasons why self-explanation prompts may be more beneficial in combination with procedural instruction.

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

Second, conceptual instruction may render self-explanation prompts unnecessary by encouraging more spontaneous self-explanation 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, Shrager, & Siegler, 1997). These metacognitive processes may be similar to self-explanation. Hence, because conceptually instructed students are not offered a correct procedure, they may engage in an unprompted sort of self-explanation to generate procedures.

There is a subtle but important distinction between the two pathways proposed above. On the first view, conceptual instruction may render explanation prompts ineffective because it helps students to 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 explanation activity required *on the part of the learner* in constructing the final mental model. Assessment of conceptual knowledge after instruction but prior to self-explanation prompts would help to test the roles these alternative pathways may play.

In contrast, self-explanation prompts continue to be helpful in combination with procedural instruction (Rittle-Johnson, 2006). First, procedural instruction seems unlikely to directly repair faulty mental models or to promote spontaneous self-explanation. Rather, 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 encourage reflection on how and when a procedure is effective. In particular, self-explanation used in combination with procedural instruction has been shown to facilitate learning correct procedures that can be adapted to solve novel transfer problems and retained over a delay (Rittle-Johnson, 2006). Children who do not explain tend to revert to using old incorrect procedures on transfer problems and after a delay. In other words, self-explanation in combination with procedural instruction strengthened and broadened correct procedures and weakened incorrect procedures, which are central components of improved procedural knowledge (Anderson, 1983).

Beyond its interaction with prior instruction, self-explanation prompts may improve learning simply because they generally increase time spent thinking about the topic. Time on task is a potentially important factor that has rarely been controlled in prior research on the self-explanation effect. The vast majority of prior self-explanation studies have held the number of examples or problems studied constant, with the result that students in the self-explanation conditions spending more time on the intervention (e.g., Atkinson, Renkl, & Merrill, 2003; Pine & Messer, 2000; Rittle-Johnson, 2006; Siegler, 1995; Siegler, 2002; Wong et al., 2002). Given that generating self-explanations generally requires much 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.

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. Thus, we equated total time spent on the intervention task by increasing the number of practice problems in the no-explain condition of Experiment 2 and found no effect for self-explanation prompts. Notably, this additional practice allowed many students to discover a new strategy. This finding raises important questions about the overall efficiency of self-explanation prompts (see also Grobe & Renkl, 2003). Only four previous experimental studies on self-explanation successfully controlled for time on task. In two, they also failed to find an effect for self-explanation (Grobe & Renkl, 2003; Mwangi & Sweller, 1998). In the other two, they did find a benefit for self-explanation while controlling for time on task (Aleven & Koedinger, 2002; de Bruin, Rikers, & Schmidt, 2007). In Aleven and Koedinger (2002), students in the no-explain condition solved additional practice problems, but those in the self-explanation condition were explicitly instructed to reference a glossary containing conceptual information and received feedback on the explanations, and this is very different from other self-explanation studies. In de Bruin and colleagues (~~2006~~), learners spent an equivalent amount of time on the same set of problems across conditions, so those in the no-explain condition did not receive addition problems to solve. It is clear that future studies should explicitly consider the amount of time on task afforded by alternative manipulations.

Conclusion

We found that conceptual instruction was more efficient than procedural instruction when both were paired with self-explanation prompts because it supported gains in both procedural and conceptual knowledge. In addition, we found that the benefits conferred by conceptual instruction may preempt the benefits conferred by self-explanation prompts, at least when controlling for time on task. All told, the data support the contention that conceptual instruction may sometimes be a more effective means for promoting learning in mathematics than procedural instruction or self-explanation prompts.

Uncited references

(Carpenter et al., 1998; Chi, 2000; Karmiloff-Smith, 1986).

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.

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.

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.

YJECP 3046 No. of Pages 21, Model 1G
26 September 2008 Disk Used ARTICLE IN PRESS

Baroody, A. J., Ginsburg, H. P., & Waxman, B. (1983). Children's use of mathematical structure. *Journal for Research in Mathematics Education, 14*, 156–168.

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.

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.

Byrnes, J. P. (1992). The conceptual basis of procedural learning. *Cognitive Development, 7*, 235–257.

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

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

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, Vol. 5: Educational design and cognitive science* (pp. 161–238). Mahwah, NJ: Lawrence Erlbaum.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students 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.

Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatley, G., Trigatti, B., et al (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.

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.

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.

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*, 197–218.

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

Grobe, 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: Lawrence Erlbaum.

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

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*, 297–312.

Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science, 14*, 511–550.

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

LeFevre, J. A., Greenham, S. L., & Waheed, N. (1993). The development of procedural and conceptual knowledge in computational estimation. *Cognition and Instruction, 11*, 95–132.

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: Externally imposed goals influence what is learned. *Journal of Educational Psychology, 92*, 734–744.

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28*, 451–466.

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.

National Mathematics Advisory Panel. (2008). Foundations for success: The final report of the National Mathematics Advisory Panel. Washington, DC: U.S. Department of Education. Retrieved July 20, 2008, from www.ed.gov/MathPanel.

Peled, I., & Segalis, B. (2005). It's not too late to conceptualize: Constructing a generalized subtraction schema by abstracting and connecting procedures. *Mathematical Thinking and Learning, 7*, 207–230.

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

Peskin, J. (1998). Constructing meaning when reading poetry: An expert–novice study. *Cognition and Instruction, 16*, 235–263.

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.

Please cite this article in press as: Matthews, P. & Rittle-Johnson, B. In pursuit of knowledge: Comparing self-explanations, ... *Journal of Experimental Child Psychology* (2008), doi:10.1016/j.jecp.2008.08.004

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. (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*, 1–29.

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., Saylor, M., & Swygert, K. (2007). Learning from explaining: Does it matter if mom is listening? Journal of Experimental Child Psychology, doi: 10.1016/j.jecp.2007.10.002.

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

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

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. Journal of Experimental Psychology:. *Learning, Memory, and Cognition, 8*, 484–494.

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.

Star, J. R., & Seifert, C. (2006). The development of flexibility in equation solving. *Contemporary Educational Psychology, 31*, 280–300.

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

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

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.