

## Promoting Transfer: Effects of Self-Explanation and Direct Instruction

Bethany Rittle-Johnson

*Vanderbilt University*

Explaining new ideas to oneself can promote transfer, but how and when such self-explanation is effective is unclear. This study evaluated whether self-explanation leads to lasting improvements in transfer success and whether it is more effective in combination with direct instruction or invention. Third- through fifth-grade children (ages 8–11;  $n = 85$ ) learned about mathematical equivalence under one of four conditions varying in (a) instruction on versus invention of a procedure and (b) self-explanation versus no explanation. Both self-explanation and instruction helped children learn and remember a correct procedure, and self-explanation promoted transfer regardless of instructional condition. Neither manipulation promoted greater improvements on an independent measure of conceptual knowledge. Microgenetic analyses provided insights into potential mechanisms underlying these effects.

Learning is often plagued by the inert knowledge problem—knowledge that is not transferred to new contexts or problems (Bransford, Brown, & Cocking, 2001). Generating explanations for oneself (i.e., self-explanation) is one promising process that can promote transfer (e.g., Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Siegler, 2002). What is less clear is whether promoting self-explanation is most effective in combination with direct instruction or discovery learning conditions and whether self-explanation is beneficial after a delay and across a range of measures (e.g., learning correct procedures, transferring the procedures, and understanding relevant concepts in the domain). This study evaluates the effects of self-explanation under different instructional conditions across a range of measures and assessment delays.

### *Discovery Learning Versus Direct Instruction*

Influential theories in psychology and reform efforts in education often claim that discovery learning supports better transfer and conceptual knowledge than direct instruction. For example, Piaget asserted in his book *To Understand Is to Invent* that “to understand is to discover, or reconstruct by rediscovery, and such conditions must be complied with if in the future individuals are to be formed who are capable of production and creativity and not simply repetition” (Piaget, 1973, p. 20). This claim is widely accepted in education communities (e.g., Bredekamp & Copple, 1997; Fuson et al., 1997; Hiebert et al., 1996; Kamii & Dominick, 1998; Stohr-Hunt, 1996; Von Glasersfeld, 1995). For example, in many reform-based mathematics programs, “a great deal of lesson time is devoted to allowing children to work out their own procedures . . .” (Fuson et al., 1997, p. 131).

In contrast, other theories (e.g., information processing theories such as cognitive load theory) propose that discovery conditions often overload working-memory capacity, and thus advocate more direct instruction (e.g., Sweller, 1988). It is well established that humans have limited working-memory capacity (Baddeley, 1992; Miller, 1956), and that well-organized knowledge structures, or schemas, are needed to overcome these working-memory limitations (Chase & Simon, 1973; Larkin, McDermott, Simon, & Simon, 1980). Direct instruction can provide organizing schemas for novices in a domain that help coordinate information in working memory (e.g.,

---

Portions of this data were presented at the 26th Annual Meeting of the Cognitive Science Society.

This research was supported by a small research grant from Peabody College, Vanderbilt University.

A special thanks to the staff, teachers and children at St. Edward School for participating in this research project. I am grateful to Jon Tapp for programming the computer task, to Jennifer Behnke and Gayathri Narasimham for collecting the data and coding explanation quality, to Kathryn Swygart, Betsy Thomas and Stephanie Crisafulli for help entering and coding the data, to Joshua Johnson for help conducting and interpreting the multiple imputations, and to Warren Lampert and David Cordray for general statistical help. Martha Alibali, Megan Saylor and Jon Star provided valuable suggestions for improving the manuscript.

Correspondence concerning this article should be addressed to Bethany Rittle-Johnson, 230 Appleton Place, Peabody #512, Vanderbilt University, Nashville, TN 37203. Electronic mail may be sent to [bethany.rittle-johnson@vanderbilt.edu](mailto:bethany.rittle-johnson@vanderbilt.edu).

Sweller, Van Merriënboer, & Paas, 1998). In contrast, discovery-learning conditions require random search, and thus are heavily constrained by capacity limitations. Sweller (2003) claimed that “Direct guided instruction, rather than problem solving, should be used as a means of acquiring schemas . . . and should always be used if available” (p. 246).

Evidence exists for the benefits of both discovery learning and direct instruction. In support of discovery learning, children who discover their own procedures often have better transfer and conceptual knowledge than children who only adopt instructed procedures (although these studies did not utilize random assignment) (e.g., Carpenter, Franke, Jacobs, Fennema, & Empson, 1998; Hiebert et al., 1996; Hiebert & Wearne, 1996; Kamii & Dominick, 1998). In support of direct instruction, a large number of studies have shown an advantage for learning and transfer if people study worked examples (a form of direct instruction) rather than solve problems unaided (a form of discovery learning) (see Atkinson, Derry, Renkl, & Wortham, 2000, and Sweller et al., 1998, for reviews). Further, students who learn via carefully designed direct instruction can demonstrate substantial transfer (Klahr & Carver, 1988; Klahr & Nigam, 2004; Schwartz & Bransford, 1998), and direct instruction on a procedure can lead to improvements in conceptual knowledge (Rittle-Johnson, Siegler, & Alibali, 2001).

This contradictory evidence may partially be explained by a confound typically found in this literature between the source of information (did someone tell you or did you discover it yourself?) and the processes that go on during acquisition (Mayer, 2004). The potential benefits of discovery learning may be due to actively engaging the learner in manipulating, linking, and evaluating information—in other words, self-explanation—rather than the discovery of the procedure itself. Successful uses of direct instruction may emerge when learners are engaged in active cognitive processes like self-explanation. How does discovery learning versus direct instruction impact learning when learners are prompted to self-explain?

#### *Effectiveness of Elicited Self-Explanation Under Conditions of Discovery or Direct Instruction*

Although there is a growing body of literature supporting the benefits of self-explanation for learning, the conditions under which it is most beneficial and the mechanisms underlying the effect are not well understood. Prior research that involved random assignment of participants to self-explain or

not can be categorized based on the instructional conditions (direct instruction or discovery) and three outcome measures: procedural learning, procedural transfer, and conceptual knowledge. Procedural learning is the ability to execute action sequences to solve familiar problems, procedural transfer is the ability to extend known procedures to novel contexts, and conceptual knowledge is understanding principles governing a domain and the interrelations between units of knowledge in a domain (e.g., Bisanz & Lefevre, 1992; Greeno, Riley, & Gelman, 1984; Rittle-Johnson et al., 2001). This brief literature review is limited to studies that focused on problem solving, rather than text comprehension, and to studies that use the most widely accepted operational definition of self-explanation, which is generating explanations of correct material by oneself (rather than explaining one’s own, potentially incorrect, solutions or adopting explanations provided by others).

First, consider prior experimental research on self-explanation using some form of direct instruction, such as studying text and/or worked examples. Several studies have found that prompts to self-explain, when compared with no-prompts, lead to immediate improvement in procedural learning (Bielaczyc, Pirolli, & Brown, 1995; Pine & Messer, 2000) as well as procedural transfer (Aleven & Koedinger, 2002; Atkinson, Renkl, & Merrill, 2003; Renkl, Stark, Gruber, & Mandl, 1998; Wong, Lawson, & Keeves, 2002). These studies have included participants ranging in age from 5 to adulthood in domains ranging from principles of balance to geometry. The transfer contexts in these studies have been highly similar to the intervention context, and only one study, by Wong et al. (2002), assessed learning after a delay (1 day to 1 week). Conceptual understanding was not assessed in any study. Other studies have failed to find a benefit for eliciting self-explanations during direct instruction in a variety of mathematical and scientific domains (Conati & Vanlehn, 2000; Didierjean & Cauzinille Marmeche, 1997; Große & Renkl, 2003; Mwangi & Sweller, 1998).

Less experimental research has been done on the effectiveness of self-explanation under discovery learning conditions. A few studies have found that learners demonstrate immediate gains in procedural learning (Curry, 2004; Neuman & Schwarz, 1998; Siegler, 1995) as well as in procedural transfer (Siegler, 2002). In these studies, learners ranging in age from 5 to adulthood solved problems and received feedback on the correct solutions before self-explaining (typically in a single session). Again, these studies did not assess conceptual knowledge and the assessment context was very similar to the

intervention context (e.g., no temporal delay). In contrast, two unpublished studies have failed to find a benefit for eliciting self-explanations under similar problem-solving conditions (Earley, 1999; Rittle-Johnson & Russo, 1999).

Self-explanation is generally accepted as an important, effective, and domain-general means to improve learning. However, a careful review of the literature reveals that prompting learners to self-explain sometimes does not improve learning and that little is known about the causal impact of self-explanation on conceptual knowledge or on knowledge improvements that persist over a delay. Further, no published study has compared the effects of self-explanation under conditions of direct instruction versus discovery learning. Establishing the conditions under which self-explanation facilitates deep learning across a variety of measures is critical to understanding and maximizing the potential of this widely used learning process. For example, under discovery learning conditions, self-explanation can promote invention of new problem-solving approaches and deeper search for more sophisticated ways of thinking (Siegler, 2002). When direct instruction on a procedure is provided, does self-explanation continue to promote invention of additional approaches, especially more sophisticated approaches, or does knowledge of a correct procedure preclude invention of additional procedures? Furthermore, both discovering procedures and spontaneous self-explanation are associated with better conceptual knowledge (e.g., Carpenter et al., 1998; Chi et al., 1989). Will elicited self-explanations promote conceptual knowledge when direct instruction is provided?

The current study evaluated whether prompts to self-explain led to improvements in procedural learning, procedural transfer, and conceptual understanding, and whether these improvements persisted over a delay. It also compared the effects of self-explanation in combination with direct instruction or discovery learning conditions. This comparison tested the hypothesis that it is not the source of information (direct instruction vs. discovery learning), but rather engagement in active processing encouraged by self-explanation, that promotes transfer and conceptual understanding. Direct instruction may discourage active processing at times, but prompts to self-explain should mitigate this and improve depth of learning.

### *This Study*

Children in this study learned under one of four conditions based on crossing two factors: (a) direct

instruction versus discovery learning and (b) prompts to self-explain versus no prompts to explain. Learning was assessed using three outcome measures (procedural learning, procedural transfer, and conceptual knowledge) immediately and after a 2-week delay. Incorporating microgenetic methods, such as the use of fine-grained and repeated knowledge assessments, allowed for a better understanding of how instructional condition and self-explanation influenced learning (Siegler & Crowley, 1991).

In this study, third- through fifth-grade students learned to solve mathematical equivalence problems, which tap the idea that two sides of an equation represent the same quantity. Mathematical equivalence is a fundamental concept in both arithmetic and algebra. Conceptual knowledge of equivalence incorporates at least three components: (a) the idea that two quantities can be equal, (b) the meaning of the equal sign as a relational symbol, and (c) the structure of equations, including the idea that there are two sides to an equation. Although fourth and fifth graders understand what it means for quantities to be equal, elementary school children often interpret the equals sign as simply an operator signal that means “adds up to” or “gets the answer,” and do not interpret it as a relational symbol meaning “the same as” (Baroody & Ginsburg, 1983; Cobb, 1987; Kieran, 1981; Rittle-Johnson & Alibali, 1999; Sfard & Linchevski, 1994). They also do not seem to understand that there are two sides to an equation or accept equations written in nonstandard formats (e.g., “ $3+4 = 4+3$ ” or “ $5 = 5$ ”) (Baroody & Gannon, 1984; Cobb, 1987; Rittle-Johnson & Alibali, 1999).

To assess conceptual knowledge, researchers typically use novel tasks for which children do not already know a solution procedure. Thus, children must rely on their knowledge of relevant concepts to generate an answer (e.g., Bisanz & Lefevre, 1992; Briars & Siegler, 1984; Greeno et al., 1984; Hiebert & Wearne, 1996). Because conceptual knowledge may be implicit, it is important to include recognition measures that do not require the verbalizing of one’s thinking (Karmiloff Smith, 1993). Tasks assessing conceptual knowledge of mathematical equivalence include defining the equal sign (i.e., an explicit knowledge measure) and categorizing equations such as “ $5 = 5$ ” as making sense or not (i.e., an implicit knowledge measure; Baroody & Gannon, 1984; Rittle-Johnson & Alibali, 1999).

Novel problems such as  $3+4+5 = 3+ \_$  challenge children’s naive understanding of equivalence in a familiar arithmetic context, and approximately 70% of fourth and fifth graders do not solve these problems correctly (Alibali, 1999; Perry, Church, & Goldin

#### 4 Rittle-Johnson

Meadow, 1988). Children can learn a correct procedure for solving the problems with some form of input (such as accuracy feedback or direct instruction), but then often do not transfer the learnt procedure to novel problem formats (Alibali, 1999; Perry et al., 1988; Rittle-Johnson & Alibali, 1999). However, prompts to self-explain under discovery learning conditions can improve procedural transfer on these problems (Siegler, 2002).

### Method

#### Participants

Initial participants were 121 third- through fifth-grade children (53 girls) from an urban, parochial school serving a working- to middle-class population. Children solving half or more of the mathematical equivalence problems correctly at pretest were excluded from the study (29% of children, mostly fourth and fifth graders; 15 girls). One additional student was excluded because he did not take the pretest. The final sample consisted of 42 third graders (16 girls), 22 fourth graders (14 girls), and 21 fifth graders (8 girls). In the final sample, approximately 10% of the participants were ethnic minorities (8% African American). According to the teachers, the children had never seen mathematical equivalence problems before and had never or rarely seen or solved problems in which the equal sign was not at the end. Teachers reported discussing the meaning of the equal sign often.

#### Design

Children participated in a pretest, intervention, immediate posttest, and delayed posttest. During the intervention, all children solved eight mathematical equivalence problems and received accuracy feedback. Children were randomly assigned to one of four conditions in the intervention based on crossing two factors: (a) instruction versus invention and (b) self-explanation versus no explanation. The number of participants in each condition was as follows: *instruct+explain* ( $n = 21$ ), *instruct+no-explain* ( $n = 22$ ), *invent+explain* condition ( $n = 22$ ), and *invent+no-explain* ( $n = 20$ ). Children from each grade were evenly distributed across the four conditions.

#### Materials

*Intervention session.* All eight 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 (e.g.,  $4+9+6=4+ \_$  and  $3+4+8= \_+8$ , which are referred to as *standard A+* and *+C* problems, respectively). The first three problems were A+ problems, and then +C problems were introduced and alternated with A+ problems to push children to generalize their procedure to both formats. The problems were presented on a laptop computer and were embedded in a game that help two astronauts get to the moon and back (see Figure 1). Immediately before and after the intervention, children solved one standard A+ and one standard +C problem (as a warm-up and a verbal-report posttest, respectively).

*Written pretest and posttests.* The same written assessment of conceptual and procedural knowledge was administered at pretest, immediate posttest, and delayed posttest, with one exception, as noted below. There were two *procedural learning* problems in standard A+ and +C format (i.e.,  $7+3+4=7+ \_$ ,

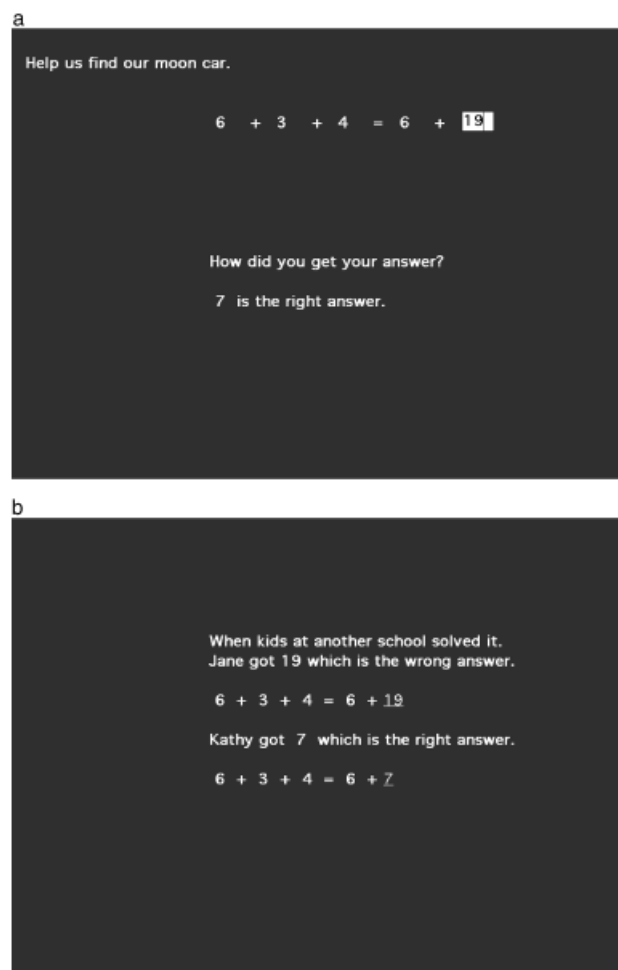


Figure 1. Screen shots from the intervention. Panel a is a completed problem and panel b is the additional screen for participants in the self-explanation conditions.

$4+5+8 = \_+8$ ). At posttest, these problem formats were familiar, and children could solve them using step-by-step solution procedures learned during the intervention. There were six *procedural transfer* problems that had no repeated addend on the right side of the equation (i.e.,  $6+4+5 = 9+\_;$   $5+7+3 = \_+4$ ), had the blank on the left side of the equation (i.e.,  $8+\_ = 8+6+4;$   $\_+9 = 6+9+3$ ), or included subtraction (i.e.,  $8+5-3 = 8+\_;$   $4+6-3 = \_-3$ ). These problem formats were unfamiliar to the children and could be solved by applying or adapting procedures learned during the intervention, which is a standard approach for measuring transfer (e.g., Atkinson et al., 2003; Chen & Klahr, 1999).

Children were encouraged to show their arithmetic calculations when solving the problems. The posttests contained all eight problems. To avoid frustrating children who were not expected to know how to solve the problems at pretest, only four of the problems were included on the pretest (two procedural learning and two no-repeated-addend transfer 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—the meaning of the equal sign and the structure of equations—and were adapted from Alibali and Grobman (2001), Baroody and Gannon (1984), and Rittle-Johnson and Alibali (1999). These items could not be solved through direct application or adaptation of a solution procedure learned during the intervention. A majority of the items were designed to tap implicit knowledge.

### Procedure

Children completed the written pretest in their classrooms in one 30-min session. Within 1 week of

the pretest, they completed a one-on-one intervention session lasting approximately 40 min, and the interval between the pretest and intervention was comparable across conditions. The session was conducted by one of two female experimenters in a quiet room at the school. During the intervention session, children first solved the two warm-up problems. After they had solved the second warm-up problem, children were told that they had solved it incorrectly to motivate them to figure out correct ways to solve the problems during the intervention.

The intervention was the same for all conditions in as many ways as possible. For each of the eight problems, all children solved the problem, reported how they solved the problem, and received accuracy feedback, including being given the correct answer. Figure 1 provides a screen shot of a completed intervention problem.

For the instruction conditions, on the first two problems (both standard A+ problems), children were taught a correct, add–subtract procedure for solving the problems. For the problem  $4+9+6 = 4+\_$ , the experimenter said: “You can add the 4 and the 9 and the 6 together before the equal sign (gesture a “circle” around the numbers), and then subtract the 4 that’s over here, and that amount goes in the blank. So, try to solve the problem using this strategy.” For the invention conditions, no instruction was given. To parallel the involvement of the experimenter on these first two problems, children in the invention conditions were prompted by the experimenter to “Think of a new way to solve the problem.” This prompt was not intended to be the primary motivator of invention. Motivation to invent was a more general feature of the context given that students received feedback that their answer was incorrect and did not receive instruction on a correct way to solve the problems.

Table 1  
Conceptual Knowledge Assessment

Concept	Task	Coding (2 points)
Meaning of equal sign	<i>Define equal sign</i>	Mention “the same” or “equal” (2 points); mention above, but limit to a specific circumstance (1 point)
	<i>Rate definitions of equal sign: rate 4 definitions as “always, sometimes, or never true”</i>	Rate “two amounts are the same” as “always true” (2 points) or “sometimes true” (1 point)
	<i>Group symbols: place symbols such as =, +, &lt;, and 5 into three groups</i>	Group =, >, and < together (2 points)
Structure of equations	<i>Recognize use of equal sign in multiple contexts: indicate whether 8 problems such as <math>8 = 2+6</math> and <math>3+2 = 6 - 1</math> make sense</i>	>75% correct (2 points); 75% correct (1 point)
	<i>Correct encoding: reproduce 4 equivalence problems from memory</i>	Correctly reproduce problem (0.5 point each)

## 6 Rittle-Johnson

Table 2  
Procedures Used to Solve Mathematical Equivalence Problems

Procedure	Sample explanation (8+7+3 = 8+_)	Frequency of use			
		Pre	Intervention	Post	Delay
Correct procedures		7	56	49	48
Equalize	I added 8 plus 7 plus 3 and I got 18 and 8 plus 10 is 18	2	7	5	9
Add-subtract	I did 8 plus 7 equals 15 plus 3 equals 18 and then 18 minus 8 equals 10	1	36	25	23
Grouping	I took out the 8's and I added 7 plus 3	1	10	9	8
Ambiguous	8 divided by 8 is 0 and 7 plus 3 is 10	3	3	10	8
Incorrect procedures		93	45	51	53
Add all	I added the 8, the 8, the 7 and the 3	41	17	10	14
Add to equal sign	8 plus 7 equals 15, plus 3 is 18	34	14	15	12
Incorrect grouping	I added 8 plus 7	0	3	4	4
Don't know	I don't know	12	3	4	4
Other	I used 8 plus 8 and then 3	6	7	17	19

The self-explanation manipulation occurred on the remaining six problems, which alternated between standard A+ and +C problems. After solving each problem and being shown the correct answer, children in the self-explanation conditions saw an additional screen with the answers that two children at another school had given: one correct and one incorrect, as shown in Figure 1. The incorrect answer matched the incorrect solution method the child had used most often at pretest—either add all or add-to-equal-sign (see the *Coding* section). The experimenter asked the participant to explain verbally both *how* the other children had obtained the answer and *why* each answer was correct or incorrect. Children in the no-explain conditions were given the correct answer, but did not see the additional screen. The intervention was audiotaped and videotaped.

Immediately after completing the intervention, children solved two standard problems and reported how they had solved each problem (*verbal-report posttest*). Then, they completed the immediate paper-and-pencil posttest, administered individually by the experimenter in the same room. Approximately 2 weeks later, the experimenters asked children to complete the delayed posttest as a group in their classrooms.

### Coding

For the procedural learning and transfer problems, the procedure children used to solve each problem was coded based on their written calculations on the assessments and on their verbal reports during the intervention (see Table 2). On the assess-

ments, procedure use could be inferred from children's written calculations; for example, if a student wrote  $7+3+4+7=21$  in her solution to  $7+3+4=7+_$ , it was coded as use of add all. 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). An accuracy score was calculated based on the percentage of problems children solved using a correct procedure, regardless of whether they made an arithmetic error. On the conceptual knowledge assessment, each item was scored from 0 to 2 points for a possible total of 10 points (see Table 1), and scores were converted to percentages. Finally, for children in the self-explanation conditions, their explanations of how and why solutions were correct and incorrect were coded. Children's "how" explanations were coded using the procedure use coding scheme (see Table 2). Table 3 lists the codes and frequencies for their "why" explanations.

Independent raters coded 20% of participants' procedure use across all phases of the study and their how/why explanations during the intervention. Interrater agreement ranged from 85% for self-explanations of why a solution was correct during the intervention to 94% for procedure use during the intervention.

### Treatment of Missing Data

Eight participants (9% of the sample) were absent from school on the day of the delayed posttest, largely because of a flu outbreak. The absent participants did not differ significantly from the nonabsent

Table 3  
Coding for Children's Explanations of Why Another Child's Answer Was Correct or Incorrect

Explanation type	Sample explanation	% trials
<i>Equal sides</i> (conceptual): recognizes two sides of an equation are equal, verbally or in gesture	They both have to equal the same thing	1
<i>Equal sign</i> (conceptual): mention equal sign or equals but doesn't mention or gesture to sides	Because there's an equal sign	6
<i>Answer</i> : refers to the quality of the answer, with no justification	Because the answer is too high	16
<i>Procedure</i> : talks about a specific procedure for solving equation	Because you need to subtract that number	27
<i>Wrong use of equal sign</i>	Equal sign should really be a minus	2
<i>Don't know</i>	I don't know	12
<i>Vague</i>	That's how you're supposed to do it	36

participants on the pretest measures. A disproportionate number of the children who were absent were from the invent+no-explain condition ( $n = 6$ ), but, as the children had no knowledge of the date of the delayed posttest, these data can be considered as missing at random. To deal with this missing data, a multiple imputations technique was used to approximate the missing accuracy scores on the delayed posttest (Rubin, 1987). The use of multiple imputations, rather than the traditional method of omitting participants with missing data, leads to more precise and unbiased conclusions (Peugh & Enders, 2004; Schafer & Graham, 2002). Simulation studies have found that using multiple imputations when data is missing at random (as in this study) leads to the same conclusions as when there is no missing data (Graham, Hofer, & Mackinnon, 1996; Schafer & Graham, 2002).

Using multiple imputations captures the uncertainty about participants' missing responses by simulating draws from the set of probable responses. To conduct the multiple imputations, 10 sets of plausible values for the missing data were created using the Markov chain Monte Carlo (MCMC) method offered in SAS and based on Schafer (1997). The imputation model included all the independent and dependent variables that were included in subsequent analyses on accuracy scores, as outlined below. Next, each of these 10 data sets was analyzed using standard complete-data methods such as ANOVA and regression. Finally, the results of the 10 analyses were combined in SAS using formulas specified by Rubin (1987) to yield a single set of results that incorporate the uncertainty in the missing data. Comparison of effect sizes for the condition manipulations indicated that the multiple imputations had minimal influence on effect-size estimates (i.e., .001 to .04 differences using Cohen's  $d$ ) compared with casewise deletion approach.

## Results

Instruction and self-explanation were both expected to influence procedural learning, procedural transfer, and conceptual understanding. To control for prior knowledge differences, pretest conceptual and procedural knowledge as well as grade level were included in all analyses as covariates. Preliminary analyses indicated that students' grade level never interacted with condition, and thus these interaction terms were not included in the final models.

### Procedural Learning

Children who were included in the study had very little knowledge of correct procedures for solving mathematical equivalence problems at pretest. Most (79%) did not solve any of the pretest problems correctly, with the remaining children solving only one of the four problems correctly. At pretest, children typically added all four numbers or added the three numbers before the equal sign (see Table 2), and there were no differences in accuracy across the different conditions.

During the intervention, children solved two types of problems, standard A+ and +C problems. Children initially solved A+ problems, and instruction was only provided on this problem type. Thus, +C problems during the intervention tested and pushed children to generalize their procedures to a second problem format. Examination of accuracy trial –by trial during the intervention (see Figure 2) suggested that instruction greatly improved accuracy on A+ problems, that children in all groups initially had difficulty generalizing their procedure to +C problems, and that self-explanation facilitated generalization to +C problems. A repeated-measures ANCOVA on accuracy scores during the intervention confirmed these observations. Problem type (A+ or +C) was a within-subject factor and

## 8 Rittle-Johnson

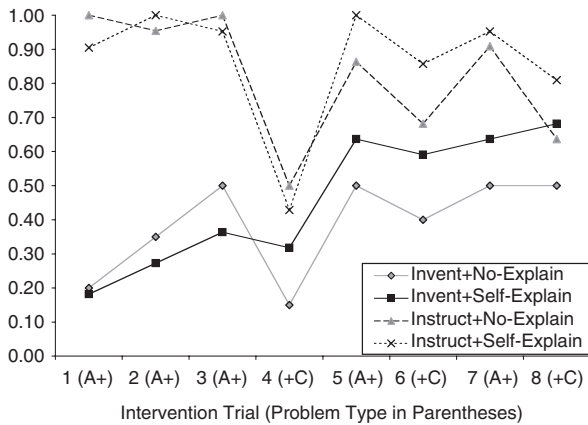


Figure 2. Proportion of children solving each intervention problem correctly, by condition.

instruction (vs. invention) and self-explanation (vs. no explanation) were between-subject factors. The control variables noted above were included to control for prior knowledge differences.

There was a main effect for problem type, such that accuracy was higher on A+ ( $M = 69\%$ ,  $SD = 40$ ) than on +C problems ( $M = 55\%$ ,  $SD = 36$ ),  $F(1, 78) = 5.44$ ,  $p = .02$ . There was also a main effect for instruction, such that accuracy was higher for children who received instruction ( $M = 84\%$ ,  $SD = 16$ ) than for those who invented ( $M = 43\%$ ,  $SD = 36$ ),  $F(1, 78) = 39.34$ ,  $p < .001$ ,  $d = 1.48$ . There was no main effect for self-explanation ( $p = .15$ ) or interaction between instruction and self-explanation ( $p = .42$ ). However, instruction, and to some extent self-explanation, interacted with problem type,  $F(1, 78) = 25.83$ ,  $p < .001$  and  $F(1, 78) = 3.52$ ,  $p = .06$ , respectively. As shown in Figure 2, instruction had a larger impact on A+ problems than on +C problems. In contrast, self-explanation predominantly influenced accuracy on +C problems. Follow-up analysis on +C problem accuracy confirmed that children who explained ( $M = 61\%$ ,  $SD = 32$ ) had greater success on +C problems than children who did not explain ( $M = 48\%$ ,  $SD = 38$ ),  $F(1, 78) = 4.27$ ,  $p = .04$ ,  $d = 0.37$ . Children who received instruction ( $M = 65\%$ ,  $SD = 32$ ) also solved more +C problems correctly compared to those who invented ( $M = 44\%$ ,  $SD = 37$ ),  $F(1, 78) = 7.81$ ,  $p = .01$ ,  $d = 0.61$ .

Improved procedural learning for the instruction and self-explanation conditions persisted on the immediate and delayed posttests. Posttest procedural learning was assessed on three occasions: on the verbal-report posttest, on the immediate paper-and-pencil posttest, and on the delayed paper-and-pencil posttest. A repeated-measures ANCOVA was conducted on accuracy scores on the procedural

learning problems, with time of assessment as a within-subject factor and instruction (vs. invention) and self-explanation (vs. no explanation) as between-subject factors.

Generating explanations and receiving instruction both led to greater procedural learning (see Table 4). Accuracy was higher for children who explained ( $M = 74\%$ ,  $SD = 34$ ) than for children who did not ( $M = 57\%$ ,  $SD = 38$ ),  $F(1, 78) = 10.18$ ,  $p = .001$ ,  $d = 0.49$ . Accuracy was also higher for children who received instruction ( $M = 75\%$ ,  $SD = 31$ ) than for children who invented procedures ( $M = 56\%$ ,  $SD = 40$ ),  $F(1, 78) = 8.53$ ,  $p = .004$ ,  $d = 0.52$ . There was no significant effect of time of assessment and no interaction between any of the factors, indicating that self-explanation and instruction led to robust learning that persisted over a delay.

### Procedural Transfer

Procedural transfer was assessed on the immediate and delayed posttests, and the six transfer problems differed from the intervention problems along several dimensions (i.e., no repetition of addends, new location of the unknown, and inclusion of subtraction). A repeated-measures ANCOVA was conducted on procedural transfer accuracy scores, with test time as a within-subject factor, and instruction (vs. invention) and self-explanation (vs. no explanation) as between-subject factors.

As shown in Table 4, there was only a main effect for self-explanation,  $F(1, 78) = 6.20$ ,  $p = .01$ ,  $d = 0.38$ . Children who were prompted to explain solved more transfer problems correctly than those who were not ( $M = 48\%$ ,  $SD = 30$  vs.  $M = 37\%$ ,  $SD = 28$ , respectively), and instruction did not influence transfer. There was no effect of test time, indicating that the benefits of self-explanation persisted on the delayed posttest.

### Conceptual Change

Children began the study with some conceptual knowledge of mathematical equivalence ( $M = 27\%$ ;  $SD = 15$ ), and pretest scores did not differ across the conditions. To help validate the measure of conceptual knowledge, conceptual knowledge scores of children who had solved at least half of the procedural knowledge problems correctly at pretest, and thus had pretested out of the study, were compared with scores of children who participated in the intervention. As expected, children who pretested out of the study had higher conceptual knowledge scores



Table 4  
Percentage Correct by Condition for the Three Knowledge Measures

Condition	Immediate posttest	SD	Delayed posttest	SD
Procedural learning				
Invent+no-explain	40	42	46	41
Invent+self-explain	68	42	64	44
Instruct+no-explain	70	40	60	47
Instruct+self-explain	79	37	75	41
Procedural transfer				
Invent+no-explain	36	40	32	29
Invent+self-explain	49	35	42	36
Instruct+no-explain	39	29	41	29
Instruct+self-explain	53	27	47	34
Conceptual knowledge				
Invent+no-explain	28	21	38	25
Invent+self-explain	22	14	30	17
Instruct+no-explain	27	18	36	18
Instruct+self-explain	29	16	33	21

( $M = 47\%$ ,  $SD = 20$ ) than children who had not,  $F(1, 116) = 34.81$ ,  $p < .001$ ,  $d = 1.21$ .

Children's conceptual knowledge was expected to improve as a result of their problem-solving experience during the intervention. By the delayed posttest, children did show gains in conceptual knowledge ( $M = 34\%$ ,  $SD = 20$ ),  $F(1, 82) = 7.24$ ,  $p = .01$ . However, these gains were not apparent on the immediate posttest ( $M = 26\%$ ,  $SD = 17$ ),  $p = .75$ .

Although children made gains in conceptual knowledge by the delayed posttest, the amount of gain was not related to instruction or self-explanation (see Table 4). Prompts to self-explain did not lead to greater increases in conceptual knowledge ( $p = .87$ ). Invention, rather than instruction, also did not lead to greater improvements ( $p = .90$ ). Rather, all groups made improvements in conceptual knowledge from their problem-solving experiences.

### Mechanisms of Change

How do instructional condition and self-explanation impact knowledge acquisition? Improved conceptual knowledge is often thought to underlie improved retention and procedural transfer, but neither self-explanation nor instruction improved children's success on the conceptual assessment. The conceptual content of children's explanations suggests why not. Children's procedure use during the intervention and on the posttests provides clues as to how self-explanation and instruction influenced learning and transfer.

*Conceptual content of self-explanations.* Children in the self-explanation conditions were prompted to

explain how other children arrived at correct and incorrect solutions and *why* the solution was correct or incorrect. This two-part question was meant to highlight that explaining why was not the same as explaining how.

Nevertheless, children rarely included any conceptual rationale in their explanations, such as referring to making the two sides equal (see Table 3). Children provided a conceptual rationale on only 7% of trials, and instructional condition (instruction vs. invention) had minimal impact (6% vs. 9% of trials, respectively). Only 36% of children ever included a conceptual rationale in at least one explanation, and these children did not show greater improvements in conceptual knowledge at posttest. Prompting children to explain correct and incorrect solutions rarely led children to think explicitly about the conceptual rationale underlying the problems, even with the inclusion of a how prompt and a why prompt.

*Procedure invention and use during the intervention.* Instruction and explanation influenced what procedures children used, how many they used, and how they discovered them. As shown in Table 5, the distribution of which procedures children used varied by condition. Not surprisingly, children in the instruction conditions almost all adopted the add-subtract procedure, and some of these children invented additional procedures. Children in the invention conditions invented a variety of correct procedures. However, almost 30% of children in the invention conditions never implemented a correct procedure. Explanation promoted invention, regardless of instructional condition (see Table 5). Indeed, during the intervention, explanation increased the number of correct procedures children used ( $M = 1.1$ ,  $SD = 0.7$  vs.  $M = 0.9$ ,  $SD = 0.5$ ),  $F(1, 78) = 4.61$ ,  $p = 0.04$ ,  $d = 0.33$ , as did instruction ( $M = 1.2$ ,  $SD = 0.4$  vs.  $M = 0.8$ ,  $SD = 0.7$ ),  $F(1, 78) = 15.81$ ,  $p < .001$ ,  $d = 0.71$ . Although a variety of procedures were used in the different conditions, only 15% of children used more than one correct procedure during the intervention, with a majority of these children being in the instruct+explain condition (see Table 5). The use of multiple correct procedures was even rarer on the posttests. Less than 5% of children used more than one correct procedure on the immediate posttest and less than 7% used more than one correct procedure on the delayed posttest, with little influence of condition.

Some children also invented overly shallow or incorrect procedures during the intervention. During the intervention, some children invented an incorrect-grouping procedure (e.g., for  $5+8+7 = \_+7$ , adding  $8+7$ ; see Table 2). Self-explanation seemed to

## 10 Rittle-Johnson

Table 5

*Percentage of Children Who Used Each Correct Procedure and Who Used At Least One or Two Correct Procedures During the Intervention, by Condition*

Condition	Correct procedure			Number of correct procedures used	
	Add–subtract	Grouping	Equalize	At least 1	At least 2
Invent+no-explain	20	30	15	60	5
Invent+self-explain	27	36	23	82	9
Instruct+no-explain	100	5	5	100	9
Instruct+self-explain	90	14	33	100	38

encourage its invention; 41% of children in the invent+explain condition used incorrect grouping, compared with 20% of children in the invent+no-explain condition. In contrast, instruction prevented invention of this procedure; none of the children who received instruction invented incorrect grouping. Children in the invent+explain group were also more likely to describe their correct use of grouping in a shallow way by not referencing the repeated addends (14% of intervention trials, compared with 7% of trials for children in the invent+no-explain condition and less than 1% of trials for children in the instruction conditions). These findings contradict those of Siegler (2002). In that study, children solved standard +C problems under discovery-learning conditions, and children in the self-explanation condition were less likely to describe a shallow version of grouping than those who did not explain. In this study, prompts to explain encouraged children to invent a range of correct and incorrect procedures.

How did children discover these new procedures? Exposure to a second problem type was one impetus for invention. As shown in Figure 2, accuracy plummeted on children's first encounter with a +C problem. Children who solved this first +C problem correctly often invented a new procedure, rather than adapting the procedure they had been using on A+ problems. The most interesting path was for children who received instruction and explained. On the first +C problem, almost 30% of the children in this condition invented a new, equalize procedure for solving the problem, rather than generalizing the instructed add–subtract procedure (compared with 5% in the instruct+no-explain condition). On their second encounter with a +C problem, fewer children used equalize (10%). Rather, children were more likely to generalize the add–subtract procedure to this new problem type. Self-explanation seemed to promote generation of a briefly used transition procedure that helped them expand the applicability of another recently learned procedure (add–subtract).

Explaining how another child arrived at the correct solution also promoted invention. Of the children in the self-explanation conditions, 33% invented at least one correct procedure while explaining how another child got the correct answer, and this was particularly true of the grouping procedure (e.g., 10/16 children who mentioned grouping first used grouping to explain how another child solved the problem correctly). Prompts to explain aided invention of new, correct procedures when children were figuring out how someone else arrived at the correct solution.

### Discussion

Prompts to self-explain led to greater learning and transfer that was maintained over a 2-week delay, regardless of instructional condition. Direct instruction, rather than invention alone, led to better procedural learning that was retained over the delay, but instructional condition had no effect on procedural transfer. Self-explanation and instruction did not interact, and neither manipulation led to greater improvements in conceptual knowledge. The use of microgenetic methods suggested potential mechanisms underlying these effects.

#### *How Self-Explanation Improves Learning and Transfer*

The current findings converge with past findings that prompting learners to generate explanations can lead to greater learning and transfer (e.g., Alevin & Koedinger, 2002; Atkinson et al., 2003; Siegler, 2002; Wong et al., 2002). These findings expand upon past research by comparing the effects of prompts to self-explain under conditions of instruction and invention, by evaluating their impact on a range of problems immediately and after a delay, and by contributing evidence for potential change mechanisms.

Prompts to self-explain promoted learning and transfer equally well under conditions of direct

instruction and discovery learning. The combination of direct instruction and self-explanation led to the greatest procedural learning, but the effects were additive. Promoting active cognitive processing via prompts to self-explain appears equally valuable under either instructional context.

Self-explanation facilitated learning correct procedures, making significant adaptations to the procedures to solve novel transfer problems, and retaining these procedures over a delay. Children who did not explain tended to revert to using old, incorrect procedures. In other words, self-explanation strengthened and broadened correct procedures and weakened incorrect procedures, which are central components of improved procedural knowledge (Anderson, 1993). Contrary to expectations, self-explanation during problem solving did not improve performance on the conceptual knowledge measure.

Combined with microgenetic analyses of learning during the intervention, these findings suggest several potential change mechanisms while failing to support other proposed mechanisms. First, self-explanation aided invention of new problem-solving approaches. Children who self-explained in the invention condition were more likely to invent at least one correct procedure, and those in the instruction condition were more likely to invent a second correct procedure, compared with children who did not self-explain. About a third of the children invented a procedure while explaining how another child found the correct answer. At other times, children invented a new procedure when confronted with a new problem format (+C problems). These new procedures were sometimes only used for a brief period (e.g., the equalize procedure), and their invention would have been missed if microgenetic methods had not been used. Similar findings of briefly used transition procedures have been reported in a microgenetic study of learning to add (Siegler & Jenkins, 1989), and these procedures seem to support invention or extension of other, more efficient procedures.

Second, self-explanation broadened the range of problems to which children accurately applied correct procedures. Children who explained were more likely to recognize that their procedure could be applied when the blank was in a different position (A+ vs. +C problems) or when non-critical features varied (e.g., using add–subtract to solve no-repeated-addend transfer problems), insights many children do not make (e.g., Rittle-Johnson & Alibali, 1999). When people learn new ideas, they often use them on an overly narrow range of problems. Recognizing the range of problems to which an approach applies, regardless of changes in surface

features, is a critical component of learning and development (Anderson, 1993; Siegler, 1996).

Third, self-explanation supported the adaptation of procedures to solve novel problems that did not allow rote application of the procedure. Solving some of the transfer problems required considerable insight into the rationale behind the procedure. For example, to apply add–subtract to transfer problems that included subtraction, children had to adapt the addition step to “combine the numbers before the equal sign using the given operations” and the subtraction step to “do the opposite of the specified operation (e.g., add if says subtract).” Self-explanation supported such flexible adaptations.

Finally, self-explanation supported retention of correct procedures over a 2-week delay. There were no effects of test time in any analysis, indicating that the benefits of self-explanation observed immediately after the intervention were maintained on the delayed posttest. Past research on self-explanation has rarely included a delay, and demonstrating transfer over a delay is an important first step in evaluating the depth of transfer supported by self-explanation. Future research should push the bar even higher and evaluate the stability of learning over changes in context that assess far transfer, such as a different knowledge domain (e.g., a balance-scale task), physical contexts (e.g., non-school setting), and temporal contexts (e.g., months later).

Some potential mechanisms identified in prior research did not seem to be central to the current study. A common characteristic of problem solving is variable procedure use and adaptive choice among procedures (Siegler, 1996). Although self-explanation did increase variability during the intervention, children rarely used multiple correct procedures at posttest, and thus adaptive choice among procedures was not a viable change mechanism. Another potential mechanism reported by Sielger (2002) was that self-explanation promoted deeper search for more sophisticated ways of thinking. However, in the current study, self-explanation did not lead to the use of more sophisticated procedures and seemed to promote invention of an incorrect, as well as a correct, procedure. Perhaps children felt freer to experiment with the generality of procedures on different problem types during the intervention (rather than the test) phase, an opportunity that did not occur in Siegler (2002).

Improved conceptual knowledge has also been proposed as a central mechanism underlying the effects of self-explanation (e.g., by generating inferences and repairing flawed or incomplete mental models) (Chi, 2000). This is based on Chi’s studies on

learning from text and descriptive analyses of the explanations produced by exemplar learners. The current study was designed to evaluate rigorously this potential mechanism in a problem-solving context, but the findings suggest that self-explanation may not have improved conceptual knowledge. Although children made gains in conceptual knowledge, prompts to self-explain did not facilitate gains on this measure. The content of children's self-explanations suggested why. Children rarely explained the rationale for why a solution was correct. Most why explanations were vague, restated whether the answer was correct or incorrect, or described the procedure for solving the problem. Prior research has also found that prompts to self-explain correct problem solutions do not lead children to generate conceptual explanations (Mwangi & Sweller, 1998; Rittle-Johnson & Russo, 1999). The failure of self-explanation to improve performance on the conceptual knowledge measure is particularly surprising given the benefit of self-explanation for procedural transfer. Why did the insights necessary to adapt the procedures to changes in problem format not translate into improved performance on unfamiliar tasks designed to assess understanding of equivalence? It may be that the insights for procedural transfer drew more on understanding of mathematical operations such as inverse operators and less on understanding of equivalence. Or, children may have made implicit gains in understanding of equivalence that were not tapped by the conceptual knowledge measure (although the measure included multiple items designed to tap implicit knowledge of the concept). It is difficult to interpret noneffects, but this is the first study of self-explanation that assessed conceptual and procedural knowledge independently, and additional research is needed to evaluate whether promoting links with conceptual knowledge is a common mechanism through which self-explanation improves procedural transfer.

#### *How Instructional Conditions Influence Learning and Transfer*

Self-explanation improved learning and transfer regardless of whether children received instruction; instructional condition influenced learning independently. As has been found in past research, direct instruction was a more reliable method for ensuring that most children learn a correct procedure (Alibali, 1999; Klahr & Carver, 1988; Klahr & Nigam, 2004; Mayer, 2004). Over a quarter of the children in the invention conditions never invented a correct procedure, even with practice and feedback. Discovery

learning relies on extensive search through the problem-solving space, a process that is very taxing on limited working-memory capacity and frequently does not lead to learning (Sweller, 1988). Instruction also seemed to prevent invention of an incorrect procedure; a substantial minority of children in the invention conditions invented an incorrect-grouping procedure, but no child who received instruction did so. At the same time, instruction did not seem to inhibit invention of additional correct procedures, particularly when children were prompted to self-explain. Overall, if you want children to use correct procedures, teaching them one is often effective and may facilitate invention of additional correct procedures.

The use of instructed procedures was also fairly robust. Children who received instruction typically generalized a correct procedure to both A+ and +C problems and were more successful on +C problems during the intervention than invention children. Children also maintained the use of these correct procedures on the delayed posttest. Although direct instruction in typical U.S. classrooms is often associated with children forgetting procedures or applying them too narrowly (Brown & Burton, 1978; Kenney & Silver, 1997), the current findings and numerous other findings suggest that, when provided with appropriate problem-solving experience, instruction can facilitate robust use of correct procedures (Klahr & Carver, 1988; Klahr & Nigam, 2004; Sweller et al., 1998).

Finally, instruction had limited influence on transfer ability. Other research has also found that instructed procedures are not transferred broadly (Alibali, 1999; Kenney & Silver, 1997; Perry, 1991; Rittle-Johnson & Alibali, 1999). However, in this study invention also did not lead to greater transfer. Even if children who failed to invent a correct procedure were omitted from the transfer analysis, the invention groups were only slightly (and nonsignificantly) more successful on the transfer problems. There was no trade-off between instruction and invention in the current study. Instruction benefited learning and retention of correct procedures, and neither instruction nor invention impacted transfer. Rather, active cognitive processing, such as self-explaining, may be the key to transfer.

#### *Limitations and Future Directions*

Implications of the results for instructional condition are constrained by the limited time and resources for invention in this study. Children had the opportunity to invent on eight problems and had no

outside resources to support invention (such as input from a peer or text). In classrooms, children may have more time to invent along with access to a variety of outside resources. Nevertheless, the findings indicate that additional research using random assignment to condition and a variety of outcome measures is needed to evaluate the trade-offs between invention and instruction.

This study focused on invention versus direct instruction on procedures, but an analogous contrast can be made for the source of explanations. Children invented their own explanations in this study, but was this essential? Several lines of research suggest that reflecting on high-quality explanations, rather than the source of the explanation (self vs. other), is key. First, there is evidence that providing children with high-quality explanations promotes learning and transfer. A key difference between higher and lower achieving countries in the TIMSS video study was how much the teacher explained connections between ideas and procedures (Hiebert et al., 2003). Experimental evidence indicates that providing children with a conceptual explanation of mathematical equivalence problems often leads children to generate and transfer a correct procedure (Perry, 1991; Rittle-Johnson & Alibali, 1999). Second, a couple of studies have manipulated whether students are provided with an explanation or need to generate their own (Crowley & Siegler, 1999; Lovett, 1992). For example, Crowley and Siegler (1999) found that the facilitative effect of explanation for transfer did not vary for children who generated correct explanations on their own and those who adopted ones provided by the experimenter. It was important that children adopted the instructed explanations; 25% of children who were told a correct explanation never verbalized this explanation. In line with the current findings, active processing promoted by explaining, rather than the initial source of the knowledge, was the key predictor of transfer.

Additional research is needed on the impact of invented versus instructed explanations, including how it interacts with invented versus instructed procedures. For example, Perry (1991) found that instruction on a conceptual explanation of mathematical equivalence only supported broad transfer when it was not given in conjunction with instruction on a procedure. In the Perry study, children did not self-explain, and the measure of transfer was categorical and required very difficult transfer. Based on the current findings and those of Crowley and Siegler (1999), I predict that direct instruction on a correct procedure and conceptual explanation for the procedure would lead to the greatest learning

and transfer if students were also prompted to self-explain.

### Conclusion

Prompts to self-explain seem to facilitate transfer equally well under conditions of invention or instruction, and these benefits persist over a delay. A growing body of research indicates that there is indeed a time for telling (Schwartz & Bransford, 1998); invention is not necessary for children to be productive and adaptive (Klahr & Nigam, 2004; Mayer, 2004; Schwartz & Bransford, 1998). What may be necessary is for people to engage in effective cognitive processes, such as generating self-explanations.

### 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.
- Alibali, M. W., & Grobman, K. (2001). *Understanding of the equal sign at the transition from arithmetic to algebra*. Unpublished manuscript.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: instructional principles from the worked examples research. *Review of Educational Research*, 70, 181–214.
- 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.
- Baddeley, A. (1992). Working memory. *Science*, 255, 556–559.
- Baroody, A. J., & Gannon, K. E. (1984). The development of the commutativity principle and economical addition strategies. *Cognition and Instruction*, 1, 321–339.
- Baroody, A. J., & Ginsburg, H. P. (1983). The effects of instruction on children's understanding of the "equals" sign. *Elementary School Journal*, 84, 199–212.
- Bielaczyc, K., Pirolli, P. L., & 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.
- Bisanz, J., & LeFevre, J. A. (1992). Understanding elementary mathematics. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 113–136). Amsterdam: Elsevier Science Publisher.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2001). *How people learn: Brain, mind, experience, and school* (expanded ed.). Washington, DC: National Academy Press.

- Bredekamp, S., & Copple, C. (1997). *Developmentally appropriate practice in early childhood programs*. Washington, DC: NAEYC.
- Briars, D., & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology*, 20, 607–618.
- Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, 2, 155–192.
- 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.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55–81.
- Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer 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 students study and use examples in learning to solve problems. *Cognitive Science*, 13, 145–182.
- Cobb, P. (1987). An investigation of young children's academic arithmetic contexts. *Educational Studies in Mathematics*, 18, 109–124.
- 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., & Siegler, R. S. (1999). Explanation and generalization in young children's strategy learning. *Child Development*, 70, 304–316.
- Curry, L. (2004). The effects of self-explanations of correct and incorrect solutions on algebra problem-solving performance. In Forbus D. Genter & T. Regier (Eds.), *Proceedings of the twenty-sixth annual conference of the cognitive science society* (p. 1548). Mahwah, NJ: Erlbaum.
- Didierjean, A., & Cauzinille Marmeche, E. (1997). Eliciting self-explanations improves problem solving: What processes are involved? *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 16, 325–351.
- Earley, C. E. (1999). Expertise acquisition in auditing: Training novice auditors to recognize cue relationships in real estate valuation. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 59(9-A), 3529.
- Fuson, K. C., Wearne, D., Hiebert, J. C., Murray, H. G., Human, P. G., Oliver, A. I., et al. (1997). Children's conceptual structures for multidigit numbers and methods of multidigit addition and subtraction. *Journal for Research in Mathematics Education*, 28, 130–162.
- 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.
- 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.
- Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K. C., Human, P., Murray, H., et al. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher*, 25, 12–21.
- Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., et al. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. (No. NCES 2003-013). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Hiebert, J., & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. *Cognition and Instruction*, 14, 251–283.
- Kamii, C., & Dominick, A. (1998). The harmful effects of algorithms in grades 1–4. In L. J. Morrow & M. J. Kenney (Eds.), *The teaching and learning of algorithms in school mathematics. 1998 yearbook* (pp. 130–140). Reston, VA: National Council of Teachers of Mathematics.
- Karmiloff, Smith A. (1993). Self-organization and cognitive change. In M. H. Johnson (Ed.), *Brain development and cognition* (pp. 592–618). Cambridge, MA: Basil Blackwell.
- Kenney, P. A., & Silver, E. A. (Eds.), (1997). *Results from the sixth mathematics assessment of the national assessment of educational progress*. Reston, Va: National Council of Teachers of Mathematics.
- Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics*, 12, 317–326.
- Klahr, D., & Carver, S. M. (1988). Cognitive objectives in a LOGO debugging curriculum: Instruction, learning, and transfer. *Cognitive Psychology*, 20, 362–404.
- Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, 15, 661–667.
- Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, 208, 1335–1342.
- Lovett, M. C. (1992). Learning by problem solving versus by examples: The benefits of generating and receiving information. In *Proceedings of the fourteenth annual conference of the cognitive science society* (pp. 956–961). Hillsdale, NJ: Erlbaum.
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, 59, 14–19.

- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.
- 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.
- Neuman, Y., & Schwarz, B. (1998). Is self-explanation while solving problems helpful? The case of analogical problem-solving. *British Journal of Educational Psychology*, 68, 15–24.
- 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.
- Piaget, J. (1973). *To understand is to invent*. New York: Grossman Publishers.
- 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.
- 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., & 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.
- Rubin, D. B. (1987). *Multiple imputations for nonresponse in surveys*. New York: John Wiley.
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.
- Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, 16, 475–522.
- Sfard, A., & Linchevski, L. (1994). The gains and the pitfalls of reification—the case of algebra. *Educational Studies in Mathematics*, 26, 191–228.
- Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology*, 28, 225–273.
- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.
- 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.
- Siegler, R. S., & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist*, 46, 606–620.
- Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.
- Stohr-Hunt, P. M. (1996). An analysis of frequency of hands-on experience and science achievement. *Journal of Research in Science Teaching*, 33, 101–109.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257–285.
- Sweller, J. (2003). Evolution of human cognitive architecture. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 215–266). San Diego, CA: Academic Press.
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251–296.
- von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. Bristol, PA: Falmer Press, Taylor & Francis.
- 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.