Promoting Flexible Problem Solving: The Effects of Direct Instruction and Self-Explaining

Bethany Rittle-Johnson (bethany.rittle-johnson@vanderbilt.edu)
Department of Psychology and Human Development, Vanderbilt University
230 Appleton Place, Peabody #512, Nashville, TN 37203

Abstract

How do people learn flexible problem-solving knowledge, rather than inert knowledge that is not applied to novel problems? Both the source of the knowledge – instructed or invented – and a central learning process – engaging in self-explanation – may influence the development of problem-solving flexibility. Seventy-seven third- through fifth-grade students learned about mathematical equivalence under one of four conditions that varied on two dimensions: 1) prompts to self-explain and 2) invention vs. instruction on a procedure. Both self-explaining and direct instruction helped students to learn a correct problem solving procedure. Self-explanation promoted transfer, whereas direct instruction had both positive and negative effects on transfer. Overall, self-explanation is an important learning mechanism underling the acquisition of flexible problem solving with or without direct instruction.

Introduction

Everyday, we are faced with new problems to solve. How do we complete our income tax return, write a new resume, or find a new route home given recent road construction? When faced with a problem repeatedly, we often develop procedures for solving the problem, i.e. step-by-step methods for solving the problem. Ideally, we learn flexible, relatively abstract procedures that we can appropriately apply to a variety of tasks so that we do not need to invent new procedures when task conditions shift. Flexible, abstract, knowledge is also a key characteristic of expertise (Chi, Feltovich, & Glaser, 1981). Thus, understanding how people develop flexible, abstract knowledge is crucial for understanding learning and development and for designing learning environments to support flexibility.

Unfortunately, people of all ages and across a large range of domains often gain inert knowledge instead – knowledge that is not applied to new situations (see Bransford, Brown, & Cocking, 2001 for a review). For example, physics students typically fail to use knowledge of physics principles, such as Newton's Laws, to solve everyday problems (Halloun & Hestenes, 1985). Indeed, even scientists sometimes fail to use their scientific knowledge to solve mundane tasks (Lewis & Linn, 1994).

How do people learn flexible knowledge, rather than simply gaining inert knowledge, and how can we support this learning? In the current study, two processes were evaluated: 1) The source of new knowledge – invention or direct instruction and 2) A potential mechanism underlying flexible learning - generating self-explanations for why and how things work.

Invention vs. Instruction

Where do new procedures come from? Typically, we invent a procedure through problem exploration or we learn a procedure from others (e.g. via imitation or direct instruction). Major theories of learning and philosophies of education differ in their emphasis on the sources of new procedures. The current paper focuses on one source of knowledge from other people – direct instruction – and compares it to inventing procedures on ones own.

Invention and learning from direct instruction can both lead to learning of the target behavior or knowledge (e.g. Judd, 1908). However, a major concern with discovery learning is that a substantial proportion of learners never invent a correct procedure or engage in correct ways of thinking (Mayer, 2004).

Another critical issue is the relative effectiveness of each source of knowledge for supporting flexible, generalizable knowledge. Direct instruction on a procedure can lead people to learn the procedure by rote, to make nonsensical errors and to be unable to transfer the procedure to solve novel problems (e.g. Brown & Burton, 1978; Hiebert & Wearne, 1986), whereas when people invent procedures, they often use the procedures flexibly in new situations (Hiebert & Wearne, 1996). Thus, there appears to be a trade-off between instruction improving problem solving on a restricted range of problems but potentially harming flexible problem solving on a broader range of problems. The current study evaluates the pros and cons of direct instruction versus encouragement to invent a procedure on a single task and evaluates the role of self-explaining as a learning mechanism under both conditions.

Self-Explaining

A potential mechanism underlying the impact of instruction and invention on procedural flexibility (and learning more generally) is learners' attempts to generate explanations for why and how things work. Successful learners typically generated explanations while studying worked-examples to problems. These explanations included identification of gaps in understanding and linkages to previous examples or sections in the text (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). Subsequent research indicates that learners ranging from 5-years-old to adulthood in domains ranging from number conservation to computer programming can learn more if they are prompted to generate self-explanations (Aleven & Koedinger, 2002; Bielaczyc, Pirolli, & Brown, 1995; Chi, de Leeuw, Chiu, & LaVancher, 1994). These findings are cooborated by findings from classroom-based research on individual differences and on cross-cultural differences in teaching practices (Stigler & Hiebert, 1997; Webb, 1991). In all cases, generating explanations is associated with greater learning. Thus, generating explanations to explain how or why things work is a critical learning process across the lifespan and across domains.

However, the informal theories or working hypotheses that learners develop are not always correct. Rather, learners often develop incorrect theories, and engaging these incorrect theories is critical to supporting learning (Bransford et al., 2001). Evidence from a variety of domains indicates that learners' incorrect informal theories are resistant to change and often persist after formal instruction that contradicts their theories (e.g. Halloun & Hestenes, 1985).

Prompting students to generate explanations for incorrect, as well as correct, solutions, beliefs, etc, may be one method for helping learners overcome incorrect prior knowledge. For example, Siegler (2002) found that prompting students to explain both correct and incorrect solutions led to greater procedural flexibility than only explaining correct solutions. In the current study, learners in the self-explanation condition were prompted to explain both correct and incorrect solutions to maximize the effectiveness of the explanation condition.

Prompting students to self-explain has been used in conjunction with a variety of sources of new information (reading a text, studying worked example, problem-solving with feedback), but has not been used in combination with direct instruction on a correct procedure nor has prior research directly compared the role of self-explaining under different sources of new knowledge. Prompting students to engage in an effective learning process after direct instruction may help students to understand and generalize the procedure. Similarly, prompts to self-explain may help students to invent and generalize correct procedures when they do not receive direct instruction (Siegler, 2002).

The current study

These issues were evaluated in the context of children learning to solve problems that tap the idea of mathematical equivalence. Mathematical equivalence is a fundamental concept in both arithmetic and algebra. Unfortunately, most children in elementary and middle school do not seem to understand equivalence, and this poses a major stumbling block for students' success in algebra (Kieran, 1981). Novel problems such as 3+4+5=3+__ challenge students' naïve understanding of equivalence in a familiar arithmetic context, and approximately 70% of fourth- and fifth-graders do not solve these problems correctly (Alibali, 1999; Rittle-Johnson & Alibali, 1999). In the current study, third through fifth graders learned to solve these mathematical equivalence problems under one of four conditions based on two factors: 1) direct instruction on a correct procedure vs. prompts to invent a new way to solve the problems and 2) prompts for self-explanations vs. no prompts.

Method

Participants. Initial participants were 121 third- through fifth- grade students from an urban, parochial school serving a working- to middle-class population. In line with previous studies using mathematical equivalence problems, 34 students (29%) solved at least half of the mathematical equivalence problems correctly at pretest and thus were excluded from the study. One student was excluded because he did not take the pretest, and 9 students were excluded because they were absent on the day of the delayed posttest, so they could not be included in the repeated-measure analyses. The final sample consisted of 37 third-graders, 22 fourth-graders, and 18 fifth-graders.

Design. Students were randomly assigned to one of four conditions based on crossing two factors: 1) instruction on a correct procedure or prompts to invent a procedure and 2) prompts to self-explain correct and incorrect solutions or no prompts. There were 20 participants in the *instruction + explain* condition, 21 students each in the *invent + explain* and *instruction-only* conditions, and 15 students in the *invent-only* condition (unequal group sizes due to random differences in absenteeism at the delayed posttest).

Procedure. Students completed the pretest in their classrooms. Students who solved at least half of the mathematical equivalence problems incorrectly participated in a one-on-one intervention session. During the intervention session, there were 3 phases: warm-up, intervention (instruction problems and practice problems, all with accuracy feedback), and follow-up. At the end of this session, students completed the immediate paper-and-pencil posttest. Approximately 2 weeks later, students completed the delayed paper-and-pencil posttest in their classrooms.

Intervention session. All problems presented during this session were in standard format (see Table 1). At the beginning of the session, students solved two warm-up problems, explained how they had solved each problem, and were told whether they had solved the second problem correctly to motivate students to try to figure out correct ways to solve the problems. During the intervention phase, all students solved 8 problems. On all problems, students explained how they solved the problem and then were told if they had solved it correctly. The first two problems were the instructional problems and were both in the format A+B+C=A+_ (*A+ problems).* For students in the instruction conditions, the experimenter explained a correct, add-subtract, procedure for solving the problem. 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 3 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." Students in the invention conditions were asked to try to figure out a new way to solve the problems.

Table 1: Procedural Knowledge Problem Types.

Problem Type	Problem Format
Standard	A+B+C=A+__ (A+) A+B+C=__+C (+C)
No repeated addend	A+B+C=D+__ A+B+C=__+D
Subtraction too	A+B-C=A+__ A+B-C=__ - C
Swap sides	A+__=A+B+C __+C=A+B+C

Next, students solved 6 practice problems that alternated between the two standard problem formats (see Table 1). After solving each of these problems, students were also told the correct answer to the problem. Students in the explain conditions were then prompted to try to explain a correct and an incorrect solution. They were shown the solution that two students at another school had gotten – one correct and one incorrect – and were asked to explain both how each student had gotten the answer and why each answer was correct or incorrect. The intervention trials were presented on a laptop computer that recorded accuracy and solution times. At the end of the intervention, students solved two follow-up problems without feedback and explained their solutions.

Assessments. The pretest, immediate posttest and delayed posttest were identical except that only a subset of the procedural knowledge problems were presented at pretest. The *procedural knowledge* assessments contained 4 types of problems, as shown in Table 1. Letters stand for numbers and indicate when a number was repeated within a problem. The standard problem formats were used during the intervention. One instance of each of the standard and no repeated addend problems was presented on the pretest. One instance of each of the problems in Table 1 were presented on the posttests. Students were encouraged to show their work when solving the problems. The 5 items on the *conceptual knowledge* assessment are shown in Table 2.

Table 2: Conceptual Knowledge Assessment Items

Item	Coding (2 pts)
Define equal sign	Mention "the same" or "equal" (2pts)
Rate definitions of equal sign: Rate 4 definitions as "always, sometimes or never true"	Rate "two amounts are the same" as "always true" (2 pts) or "sometimes true" (1 pt)
Group Symbols: Place symbols such as =, +, <, & 5 into three groups	Group =, >, and < together (2 pts)
Recognize use of equal sign in multiple contexts: Indicate whether 8 problems such as 8=2+6 and 3+2=6-1 make sense	7 or 8 correct (2 pts); 6 correct (1 pt)
Correct Encoding: Reproduce 4 equivalence problems from memory	Correctly reproduce problem (.5 point each)

Table 3: Procedures for Solving Equivalence Problems

Procedures	Sample student explanation
Correct Procedures	
Equalize	"I added 8 plus 7 plus 3 and I got 18 and 8 plus 10 is 18."
Add-subtract	"I did 8 plus 7 equals 15 plus 3 equals 18 and then 18 minus 8 equals 10"
Grouping	"I took out the 8's and I added 7+3."
Incorrect Procedures	
Add all	"I added 8+7+3+8, which is 26"
Add to equal sign	"8 plus 7equals15, plus 3 is18."
Incorrect Grouping	"I added 8 plus 7."

Coding. On the procedural knowledge assessments, students' percent correct was used (arithmetic slips were ignored). Students' verbal explanations during the intervention were used to code students' procedure use on those problems (see Table 3). On the conceptual knowledge assessment, each item was scored from 0-2 points for a possible total of 10 points (see Table 2).

Results

The effects of condition were assessed on three outcomes: procedural learning, procedural transfer, and conceptual knowledge. Procedural learning was assessed 3 times: verbally at the end of the intervention and on the immediate and delayed posttests. Procedural transfer and conceptual knowledge were assessed twice – on the immediate and delayed posttest. Results were evaluated for each outcome using repeated-measures ANOVAs with time of assessment as a within-subject factor and instruction vs. invention and prompts to explain (yes/no) as between subject factors. Pretest conceptual and procedural knowledge were included in all analyses as covariates.

Procedural Learning

First, consider procedural learning, which was assessed using problems identical in form to those presented during the intervention (see Table 1). As shown in Figure 1, generating explanations and, to some extent, receiving instruction led to greater accuracy on the learning problems. There was a main effect for explaining, $F(1, 71) = 6.11$, $p = .02$, $\eta_p^2 = .08$, and a marginal effect for instruction, $F(1,71) = 2.84$, $p = .10$, $\eta_p^2 = .04$, and no interaction between the two conditions and no effects of time of assessment.

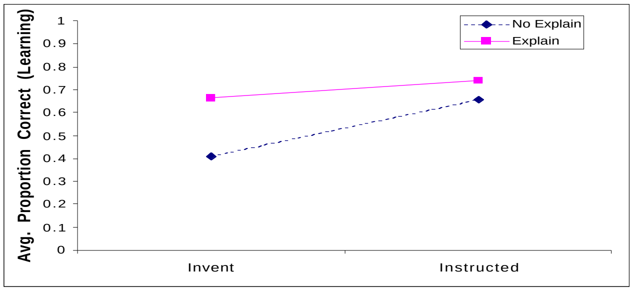

Figure 1: Effect of condition on learning problems

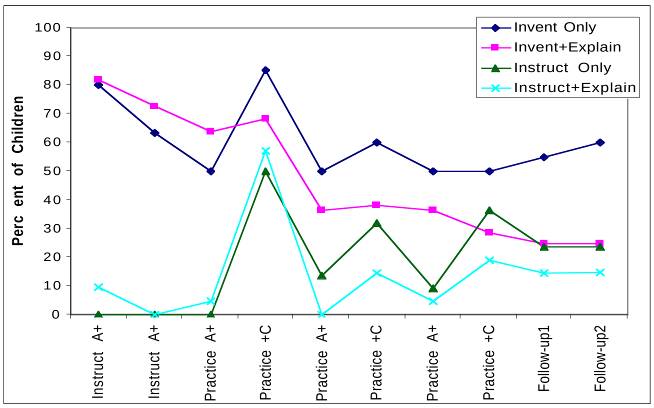

Figure 2: Incorrect procedures: Trial-by-trial use during the intervention and follow-up, by condition.

Inspection of students' procedure use during the intervention provided insights into learning pathways (see Table 3 for a description of each procedure). First, consider students' use of incorrect procedures. As shown in Figure 2, during the instruction phase (first two problems), students in the instruction condition quickly abandoned their incorrect procedures whereas most students in the invention conditions persisted in using incorrect procedures. During the practice phase, when students first encountered a problem in a different format (+C problem), there was a sharp return to using incorrect procedures. Students in the instruction condition continued to struggle with the +C problems, especially if not prompted to self-explain. Students in the invent-only condition struggled across problems – at least 50% continued to use incorrect procedures, whereas students who were prompted to self-explain steadily decreased in use of incorrect procedures.

Next, consider students' use of the correct instructed procedure – add-subtract (see Figure 3). Students in the instruction conditions quickly learned the add-subtract procedure and many students persisted in using this procedure across a majority of problems. However, a third of the students did not apply the procedure when the surface structure of the problem changed (+C problems), even though the procedure required only a very minor adaptation. Of students in the invent conditions, about 15-20% invented and used this procedure. Next, consider the other commonly used correct procedure – grouping (see Figure 4). Students in the invent conditions gradually increased their use of this procedure. Prompts to explain also helped

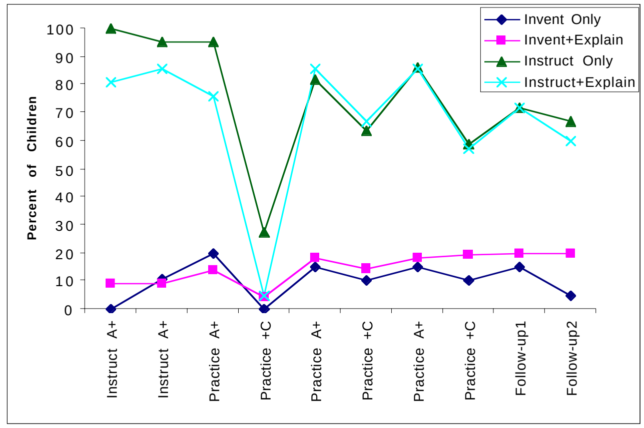

Figure 3: Use of Add-Subtract procedure trial-by-trial during the intervention and follow-up, by condition.

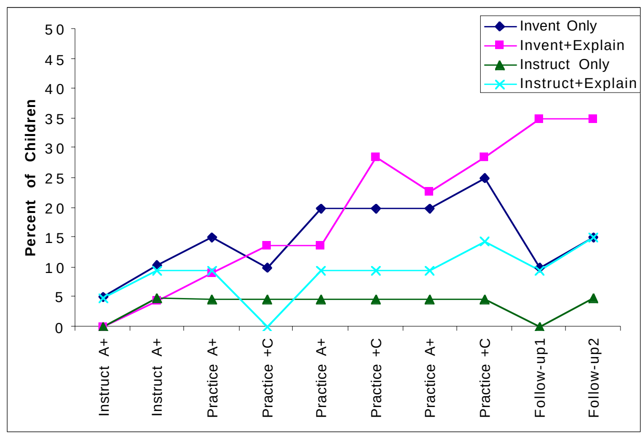

Figure 4: Use of Grouping procedure trial-by-trial during the intervention and follow-up, by condition.

students to invent and maintain use of the procedure on the follow-up problems. Finally, students in the invent conditions also used the equalizer procedure on 12% of intervention trials, whereas students in the instruct+explain condition used it on 6% of trials and students in the instruct-only condition used it on less than 1% of trials.

Overall, over half of students in the invent-only group did not learn a correct procedure through feedback alone. Less than a quarter of students in the other conditions had similar difficulty. Rather, direct instruction quickly led children to adopt a correct procedure, and prompts to explain helped students to invent new procedures.

Procedural Transfer

Next consider students' ability to transfer their procedures to novel problems (see Figure 5). Overall, there was a main effect of explaining $F(1, 71) = 3.93$, $\underline{p} = .05$, $\eta_p^2 = .05$ and no overall effect of instruction, interaction between the two, or effect of test time. Prompts to explain supported transfer of procedures to novel problems, but instruction vs. invention did not have a general effect on transfer (although see below for an important caveat). Inspection of success on individual problems suggested that the impact of

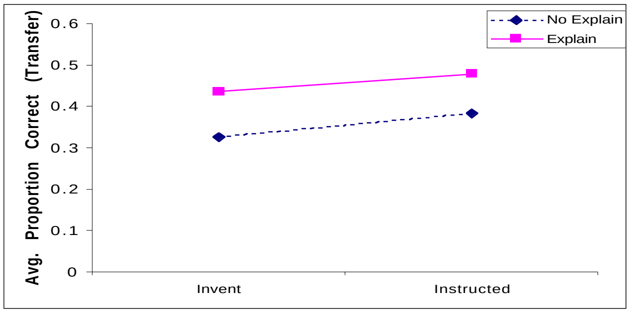

Figure 5: Effect of condition on transfer performance

condition varied by problem type. To evaluate this, a second repeated-measures ANOVA was conducted with transfer problem type (no repeated addend, including subtraction, swapping sides; see Table 1) as an additional within-subject factor. Indeed, there was an interaction between problem type and instruction, F (2, 142) = 9.37, $p < .001$, $\eta_p^2 = .12$, but no interaction of explaining with problem type. Follow-up analyses indicated that instruction improved performance on problems without a repeated addend, F (1, 71) = 7.89, $p = .006$, $\eta_p^2 = .10$. The instructed procedure did not need to be modified to solve problems without a repeated addend. Instruction had no reliable effect on the other problem types. However, focusing on the most difficult individual problem (A+B-C=_-C), receiving instruction actually harmed performance, regardless of explaining, F (1,71) = 13.12, $p = .001$, $\eta_p^2 = .16$. For example, on the delayed posttest, very few of the students who received instruction solved the problem correctly, whereas at least a third of students in the invent conditions solved the problem correctly (see Figure 6). This may be because the invented grouping procedure is easier to apply to this problem than add-subtract.

Conceptual Improvement

Finally, there was no effect of condition on gains in conceptual knowledge. Although students as a whole made small gains in conceptual knowledge from pretest to immediate posttest (m = 2.9 vs. 3.1 out of 10), $t(76) = 2.0$, $p = .05$, and made even great gains after a delay (m = 3.8), $t(76) = 5.7$, $p < .001$, the amount of gain did not vary by condition.

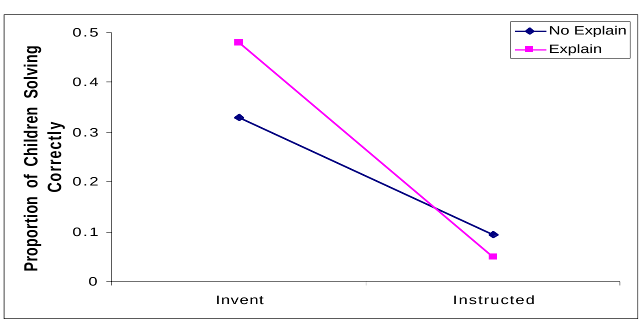

Figure 6: Effect of condition on hardest transfer problem, A+B-C=_-C, at delayed posttest

Discussion

Self-explanation is a critical learning mechanism that leads to greater procedural flexibility. The current findings converge with past findings that better learners spontaneously produce self-explanations and that prompting learners to generate explanations leads to greater learning (Aleven & Koedinger, 2002; Bielaczyc et al., 1995; Chi et al., 1989; Chi et al., 1994). The current findings expand past research by demonstrating that self-explanation is an important learning mechanism regardless of instruction. For students who received direct instruction on a correct procedure, prompts to self-explain had little influence on the use of the instructed procedure. Rather, the prompts promoted generation of additional correct procedures. Indeed, 57% of students in this condition used at least two correct procedures during the intervention, compared to only 24% of students who received instruction but were not prompted to explain. Using multiple procedures is a common feature of development and is beneficial to performance (Siegler, 2002). Prompts to explain under invention conditions also promoted invention of a correct procedure (Siegler, 2002).

Overall, students in the explain conditions were better able to solve transfer problems, regardless of instruction. Analysis of students' explanations revealed that students rarely explained the rationale for why a solution was correct. Approximately 8% of explanations included mention of equal sides or the importance of the equal sign. Rather, most why explanations where ambiguous or described the procedure for solving the problem. Combined with the finding that explanations did not influence conceptual learning, this suggests that prompts to self-explain on a problem-solving task promote exploration of alternative procedures but not reflection on conceptual-underpinnings of the procedure. Self-explanations are one promising mechanism for explaining why some learners make improvements in conceptual understanding after learning a new procedure while others do not (Rittle-Johnson & Alibali, 1999), but the current study does not support this hypothesis.

The current findings have important implications for the debate between use of direct instruction vs. encouragement to invent procedures. There are serious limitations to relying on people inventing correct procedures without guidance on effect learning processes. Half of the students in the current study never invented a correct procedure when receiving feedback on the correct answers alone. Some prior research has suggested that feedback is critical to supporting invention during exploration, but even this level of support was insufficient for many learners (e.g Lacher, 1983). In comparison, as in previous studies, direct instruction supported rapid adoption of a narrowly used procedure (e.g. Alibali, 1999). A third of the students in the instruction groups failed to generalize the add-subtract procedure even when receiving feedback during the intervention, and instruction only supported transfer to a very similar problem that required no adaptation to the procedure. On the hardest problem, having received instruction interfered with problem solving. Overall, direct instruction by itself appears to be a quick route to inert knowledge. However, promoting engagement in effective learning processes, such as self-explaining, helped students to avoid many of the downsides of both invention and direct instruction.

Overall, it is not the source of procedure, but rather engagement in a fundamental learning process, self-explanation, that is important for promoting flexible problem-solving.

Acknowledgments

This research was supported by a small research grant from Peabody College, Vanderbilt University. A special thanks to Jennifer Behnke and Gayathri Narasimham for collecting the data and to Kathryn Swygart, Betsy Thomas and Stephanie Crisafulli for help coding and entering the data.

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*(2), 147-179.

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35*(1), 127-145.

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*(2), 221-252.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2001). *How People Learn: Brain, Mind, Experience, and School* (Expanded Edition ed.). Washington, D.C.: National Academy Press.

Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science, 2*(2), 155-192.

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*(2), 145-182.

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439-477.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

Halloun, I. A., & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics, 53*, 1045-1055.

Hiebert, J., & Wearne, D. (1986). Procedures over concepts: The acquisition of decimal number knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 199-223). Hillsdale, NJ: Erlbaum.

Hiebert, J., & Wearne, D. (1996). Instruction, Understanding, and Skill in Multidigit Addition and Subtraction. *Cognition and Instruction, 14*(3), 251-283.

Judd, C. H. (1908). The relation of special training to general intelligence. *Educational Review, 36*, 28-42.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12*(3), 317-326.

Lacher, M. B. (1983). Effects of feedback, instruction, and initial performance level upon training and persistence of verbal rehearsal. *Journal of General Psychology, 108*(1), 43-54.

Lewis, E. L., & Linn, M. C. (1994). Heat energy and temperature concepts of adolescents, adults, and experts: Implications for curricular improvements. *Journal of Research in Science Teaching, 31*(6), 657-677.

Mayer, R. E. (2004). Should there be a thre-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59*(1), 14-19.

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*(1), 175-189.

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). Cambrdige: Cambridge University Press.

Stigler, J. W., & Hiebert, J. (1997). Understanding and Improving Classroom Mathematics Instruction. *Phi Delta Kappan, 79*(1), 14-21.

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education, 22*(5), 366-389.