Available online at www.sciencedirect.com
SCIENCE DIRECT®
Contemporary Educational Psychology 31 (2006) 280–300
Contemporary Educational Psychology
www.elsevier.com/locate/cedpsych

The development of flexibility in equation solving

Jon R. Star [a,*], Colleen Seifert [b]

[a] *Department of Counseling, Educational Psychology, and Special Education, 513 Erickson Hall, Michigan State University, East Lansing, MI 48824, USA*
[b] *Department of Psychology, University of Michigan, USA*

Available online 23 September 2005

Abstract

This paper explores the development of students' knowledge of mathematical procedures. Students' tendency to develop rote knowledge of procedures has been widely commented on. An alternative, more flexible endpoint for the development of procedural knowledge is explored here, where students choose to deviate from established solving patterns on particular problems for greater efficiency. Students with no prior knowledge of formal linear equation solving techniques were taught the basic transformations of this domain. After instruction, students engaged in problem-solving sessions in two conditions. Treatment students completed the "alternative ordering task," where they were asked to re-solve a previously completed problem but using a different ordering of transformations. Those completing alternative ordering tasks demonstrated greater flexibility.

Keywords: Procedural learning; Flexibility; Linear equation solving; Mathematical problem solving

1. Introduction

For much of this century, mathematics educators have sought to address students' tendency to view school mathematics as a series of procedures to be memorized. Researchers in mathematics education concur that (a) procedures learned by rote

* Corresponding author. Fax: +1 517 353 6393.
E-mail address: jonstar@msu.edu (J.R. Star).

0361-476X/$ - see front matter
doi:10.1016/j.cedpsych.2005.08.001

are easily forgotten, error-prone, and resistant to transfer; and (b) the learning of procedures must be connected with conceptual knowledge to foster the development of understanding (e.g., Hiebert & Carpenter, 1992). The National Council of Teachers of Mathematics has articulated this emphasis on conceptual learning by calling for decreased attention to "memorizing rules and algorithms; practicing tedious paper-and-pencil computations; memorizing procedures...without understanding" (National Council of Teachers of Mathematics, 1989, p. 71), and "rote memorization of facts and procedures" (National Council of Teachers of Mathematics, 1989, p. 129).

There is little doubt that the rote execution of memorized procedures does *not* constitute mathematical understanding. However, there are other ways in which a procedure can be executed other than by rote. Of particular interest in the present paper is the development of *flexibility* in the use of mathematical procedures.

2. Defining flexibility

Colloquially, flexibility refers to the ability to change according to particular circumstances. A more formal definition of flexibility will be presented below; however, it is useful to begin with a series of examples to illustrate what is meant by this construct. Table 1 shows a set of example problems, with solutions emblematic of more and less flexible solvers provided.

The following features can be seen in the solutions provided in Table 1. First, although both solvers solve all three equations correctly, the solutions in Table 1 indicate that the more flexible solver completes all three example problems using three different solution procedures, while the less flexible solver uses the same algorithm on all problems. The algorithm that the less flexible solver uses on all three

Table 1
Example problems and solutions emblematic of more and less flexible solvers

	Solutions emblematic of a *more flexible* solver	Solutions emblematic of a *less flexible* solver
$4(x+1)=8$	$4(x+1)=8$ $x+1=2$ $x=1$	$4(x+1)=8$ $4x+4=8$ $4x=4$ $x=1$
$4(x+1)+2(x+1)=12$	$4(x+1)+2(x+1)=12$ $6(x+1)=12$ $x+1=2$ $x=1$	$4(x+1)+2(x+1)=12$ $4x+4+2x+2=12$ $6x+6=12$ $6x=6$ $x=1$
$4(x+1)+3x+7=8+3x+7$	$4(x+1)+3x+7=8+3x+7$ $4(x+1)=8$ $x+1=2$ $x=1$	$4(x+1)+3x+7=8+3x+7$ $4x+4+3x+7=3x+15$ $7x+11=3x+15$ $4x=4$ $x=1$

problems will be referred to as the standard algorithm[1] for this domain. This standard algorithm has the following steps:

1. Use the distributive property to "expand" the parentheses (EXPAND[2]).
2. Transform the equation to a standard form ($ax + b = cx + d$) by combining all the variable terms and constant terms on each side (COMBINE).
3. Get the variable terms to the left side and the constants to the right side (SUBTRACT FROM BOTH).
4. Divide by the coefficient of the variable term on the left side (DIVIDE).

It is certainly the case that successful solvers in a domain have knowledge of standard algorithms; at issue here is why some solvers choose to deviate from the standard algorithm under certain conditions, which will be considered as one indicator of flexibility.

Second, note that the solutions of the more flexible solver are more efficient on all three example problems, where a more efficient solution is one that requires the application of fewer transformations to reach the solution. This increased efficiency results from a more clever or innovative application of equation solving transformations in particular ways on certain problems. On the first problem, the more flexible solver chooses to divide both sides by 4 as a first step, before distributing, despite the fact that the standard algorithm calls for this transformation to be applied always as a final step, after distributing. On the second problem, the more flexible solver temporarily changes the variable, from x to $(x + 1)$, to combine the $(x + 1)$ terms, although the standard algorithm calls for distributing at this point in the solution. And on the third problem, the more flexible solver recognizes the presence of common $(3x + 7)$ terms on both sides of the equation and in essence "cancels" this expression from both sides as a first step, which is again a deviation from the standard algorithm that results in a more efficient solution.

As these examples are intended to illustrate, we define a flexible solver as one who (a) has knowledge of multiple solution procedures, and (b) has the capacity to invent or innovate to create new procedures. With respect to (a), it is perhaps obvious that a flexible solver should have broad knowledge of procedures in a domain. An inflexible solver must rely on a small set of procedures, because this is all that she knows how to do. A flexible solver has more expertise in the domain and thus has a greater range of problem-solving strategies from which to choose (Gick, 1986; Krutetskii, 1976; Sweller, 1988; Sweller, Mawer, & Ward, 1983). Metaphorically, by having knowledge of multiple solution procedures, flexible solvers have more "tools" in their

[1] This solution procedure is "standard" in that it is commonly (and often explicitly) taught in US schools as the best way to solve most linear equations.

[2] For the remainder of this paper, each type of equation solving transformation will be referred to with a word or phrase written in all capital letters. The possible transformations are: use the distributive property (EXPAND), combine like variable terms (COMBINE), combine like constant terms (COMPUTE), move variable or constant terms to other side (SUBTRACT FROM BOTH), and multiply or divide to both sides (DIVIDE).

procedural "toolbox." For assessment purposes, this aspect of flexibility can be seen by looking at a collection of problems and determining whether a student uses several, or only a few, solution procedures on these problems. Note that the collection of problems on which knowledge of multiple solution procedures can be assessed must exhibit some of the variation that exists in the domain to enable flexible solvers to demonstrate their greater knowledge of solving strategies.

While knowledge of multiple solution procedures is clearly necessary, it does not appear to be sufficient for flexibility. Not only does a flexible solver have knowledge of multiple solution procedures, but she also has the capacity to invent new procedures for solving unfamiliar problems or when seeking an optimal solution for familiar problems. The capacity for invention of procedures has been widely studied, particularly in arithmetic (Blöte, Klein, & Beishuizen, 2000; Blöte, Van der Burg, & Klein, 2001; Carroll, 2000; Hiebert & Wearne, 1996; Resnick, 1980; Resnick & Ford, 1981; Siegler, 1996; Siegler & Jenkins, 1989). Students frequently invent their own procedures for solving arithmetic problems (Resnick, 1980; Resnick & Ford, 1981), and those who invent have been found to be more flexible solvers and to have greater conceptual understanding (Blöte et al., 2001; Carpenter, Franke, Jacobs, Fennema, & Empson, 1998; Carroll, 2000). At its core, invention in the use of algebra procedures involves the use of transformations in unusual, atypical, and innovative combinations. Metaphorically, being capable of invention means that a flexible solver has the ability to use the "tools" in her toolbox in non-standard ways that do a better job of performing certain kinds of tasks.

With respect to assessing the capacity for invention, two important clarifications are relevant. First, the capacity for invention has been defined here at the transformation level, rather than at the whole procedure level. Referring back to Table 1, the more flexible student's atypical use of the COMBINE transformation in a particular problem (combining $(x + 1)$ terms before distributing) is considered an invention, rather than the multi-step procedure which includes this atypical step. The decision to define invention at the transformation level is made for two reasons. First, prior work on flexibility in the use of algebra procedures (e.g., Lewis, 1981) suggests that whole procedure invention is quite rare. Using a more fine-grained measure of flexibility (at the transformation level) allows for a more careful investigation of the development of this capacity. And second, whole procedure invention typically involves the use of several transformations in atypical ways. Consider the more flexible student's work on the second problem in Table 1. Her solution is more efficient because she invents a new way to use the COMBINE transformation *and* because she divides both sides by 6 as a next step (which is itself an invention; had she followed the standard solution procedure from the step $6(x + 1) = 12$ by expanding first, her solution would not have been more efficient). By defining invention at the transformation level, it becomes more possible to capture the smaller steps that appear to lead to the generation of complete, maximally efficient procedures.

Second, invention, as defined here, is *not* intended to capture the idiosyncratic, inefficient, and sometimes strange solution methods that many students discover and use. It has been frequently shown that students can come up with surprising and unusual (but not particularly efficient or clever) solutions to a broad range of problems (Resnick & Ford, 1981). Rather, the capacity for invention is reserved for those students who come up with ways of using transformations atypically that have the potential to do a better job of solving certain kinds of problems.

3. The development of flexibility

How does flexibility in the use of algebra procedures develop? One hypothesis comes from studies where participants were asked to solve a problem repeatedly to observe changes in their solutions that emerged with practice. There is ample evidence that solving a problem multiple times can lead to more rote execution (Anzai & Simon, 1979; Blöte et al., 2000; Simon & Reed, 1976) and/or set effects (e.g., Luchins, 1942). However, there is also reason to hypothesize that, under certain conditions, re-solving previously completed problems can lead to more flexible solving behavior (Blöte et al., 2000; Blöte et al., 2001; Klein, 1998).

In the present study, this hypothesis is tested by utilizing a task referred to as the "alternative ordering task" (Star, 2001/2002). Participants are asked to re-solve previously completed problems but using a different ordering of transformations. In this task, students are not merely practicing the same solution over and over again, but instead are generating, comparing, and evaluating the effectiveness and efficiency of different solution procedures.[3]

There are at least two reasons to speculate that tasks such as the alternative ordering task may lead to greater flexibility. First, asking students to re-solve previously completed problems has the effect of reducing the specificity of the problem-solving goal: Rather than a specific goal of finding the particular solution to a problem, students instead complete the more general task of completing the problem using one of many different solution strategies. Sweller and colleagues (Mawer & Sweller, 1982; Sweller, 1983; Sweller et al., 1983) have demonstrated that, by reducing the specificity of the goal, novices can be guided toward more expert strategies.

Second, the alternative ordering task shares two key features with self-explanation (defined as the process by which learners generate their own explanations to justify the steps in a worked example), which have been found to facilitate learning and transfer under certain conditions (Chi, 2000; Chi & Bassok, 1989; Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, DeLeeuw, Chiu, & LaVancher, 1994; Chi & VanLehn, 1991; Pirolli & Recker, 1994; Renkl, 1997; Renkl, Stark, Gruber, & Mandl, 1998). First, a self-explanation was found to be especially beneficial to learning when it mentioned[4] a purpose or goal for a particular action (Chi et al., 1989; Renkl, 1997; Renkl et al., 1998). Second, self-explanations that display anticipative reasoning are

[3] The alternative ordering task is similar to what Klein has called the "flexibility-on-demand" task or FDT (Klein, 1998). The FDT was used by Klein as a way to assess students' flexibility (Blöte et al., 2001; Klein, 1998); arithmetic learners were asked to solve a problem two times in a row using two different procedures. If a student produced two different solution procedures, this was taken as an indication of her flexibility.

[4] Note that the act of self-explaining may or may not be spoken out loud (Chi, 2000).

more strongly linked to learning (Reimann, Schult, & Wichmann, 1993; Renkl, 1997). Anticipative reasoning refers to the generating of solutions or answers to problems or questions before such solutions are presented in the problem or text. Self-explanations that include anticipatory reasoning show that a learner is thinking and working ahead, providing the opportunity to compare the anticipated steps with the ones presented subsequently in the problem or text.

These two features of self-explanations are also present in the alternative ordering task. As students re-solve previously competed problems, they may come to know the various goals that are applicable for a particular transformation on certain problems and may subsequently identify and repair gaps in their existing knowledge of the domain, perhaps leading to greater flexibility. Furthermore, as students become familiar with the alternative ordering task, they are aware that they will be asked to re-solve the same equation again, using a different ordering of steps. As a result, it is possible that students' initial equation-solving strategies will be selected and implemented with an eye toward how the same equation would be approached the second time (e.g., anticipatory reasoning).

The goals of this study were to explore the development of flexibility in the use of algebra procedures and to explore the instructional conditions that facilitate the emergence of this outcome.

4. Method

4.1. Participants

The 36 sixth grade participants (20 female, 16 male) in this study were recruited on a first-come, first-served basis using flyers that were distributed late in the school year to all sixth grade students in the public school district of a medium-sized city in the Midwestern United States. Participants were paid \$50, in the form of a gift certificate to a local bookstore, to participate in this research.

The participants attended schools with an integrated middle school mathematics curriculum; symbolic algebra is not introduced in this curriculum until after the sixth grade, and it is not covered at all in the K-5 curriculum. Thus, it is unlikely that, at the time of the study, any of the participants had received any formal instruction in the use of symbolic methods of equation solving. However, a pre-test was administered to assess students' prior knowledge on the first day of the study (see Table 2). The problems on the pre-test were selected to determine whether students had any knowledge of formal equation solving procedures.

4.2. Procedure

Participants attended 1-h experimental sessions for five consecutive days (Monday to Friday), at the same time each day. Students met in groups of six. The first experimental session on Monday was devoted to administering a pre-test, providing instruction (described below), and administering a post-instruction test. The final

Table 2
Equations attempted on the pre-test and post-test

#	*Problem*
1	$x + 7 = 10$
2	$2x = 18$
3	$4x + 6 = 22$
4	$2(x + 5) = 22$
5	$4x + 3x = 21$
6	$6(x + 4) = 3(x + 4) + 6$
7	$2x + 4x = 12 + 6$
8	$15x + 10 = 5x + 20$
9	$2x + 4 + 8 = 20$
10	$4(x + 3) = 16x$
11	$5(x + 3) + 10x = 35 + 5x$
12	$3(x + 1) = 6(x + 1)$
13	$0.3x + 0.2 = 1.1$
14	$5(x + 1) + 10(x + 1) + 5(x + 1) = 5x + 10x$
15	$3(x + 1) + 6(x + 1) + 6x + 9 = 6x + 9$
16	$4(x + 2) + 6x + 10 = 2(x + 2) + 8(x + 2) + 6x + 4x + 8$
17	$4x + 3x + 5x + 4 = 3x + 5x + 16$
18	$3(x + 2) + 9(x + 2) = 6(x + 2)$
19	$2(x + 3) + 4x + 8 = 4(x + 2) + 6x + 2x$

session on Friday was used to administer the post-test (which was the same as the pre-test). During the three remaining sessions (Tuesday to Thursday), students solved algebra equations for the entire hour (described in depth below). All sessions took place in a seminar room on the campus of a local university.

In the first session, students were given a scripted, 30-min lesson on the transformations of equation solving. In this lesson, students were introduced to five basic transformations used in solving linear equations: combining constants (COMPUTE), combining like variable terms (COMBINE), using the distribute property (EXPAND), adding/subtracting a constant or variable term from both sides (SUBTRACT FROM BOTH), and multiplying/dividing to both sides (DIVIDE). In this initial lesson, students were never given any strategic instruction as to how the transformations could be used together to solve equations. The focus on instruction was strictly on pattern recognition: identifying which transformation could be used for particular patterns of symbols, and how that transformation was correctly applied. The intent in this brief period of instruction was to provide novice algebra learners with sufficient or prerequisite (Zhu & Simon, 1987) knowledge to enable them to begin to solve very straightforward equations; no worked-out examples were presented during instruction. Immediately following the lesson, students were given a post-instruction test, which was intended to evaluate whether they had learned the material provided during instruction. The post-instruction test assessed whether students could individually apply each of the five equation-solving transformations.

At the conclusion of the first instructional session, the groups of six in each study were randomly assigned to either a control group or a treatment group for the remainder of the week. Eight males and 10 females were randomly assigned to the control group, and the remaining 8 males and 10 females were assigned to the treatment group. The treatment and control groups differed only in that the treatment group completed alternative ordering tasks, while the control group did not. On alternative ordering tasks, students were given problems that they had previously solved and asked to re-solve them, but using a different ordering of transformations. On problems where the treatment group was asked to provide an alternative solution, the control group completed a different but isomorphic problem. For example, both treatment and control groups were asked to solve the problem, $4x + 10 = 2x + 16$. Students in the treatment group were given this same problem again and were asked to solve it using a different ordering of transformations. The students in the control group were given a structurally equivalent problem, $6x + 9 = 3x + 12$, instead. Treatment group students regularly engaged in the alternative ordering task; of the 45 problems that students could have attempted, treatment students were asked to resolve a previously solved problem 18 times. In other words, problems relating to the use of the alternative ordering task (including both the first and second attempt at each problem) accounted for 36 of the 45 problems. (See Appendix A for a list of all problems solved by students during the problem-solving sessions.) Prior to completing the alternative ordering task for the first time on Tuesday, treatment students were introduced to this task with an example. The example involved telling a robot how to make a peanut butter and jelly sandwich, and then telling the robot how to make the same sandwich but using a different ordering of steps.

During the three problem-solving sessions (Tuesday, Wednesday, and Thursday) students worked individually on all problems. Students sat alone at tables and were positioned far enough from each other so that it was impossible for any participant to see the work of another. If a student became stuck while attempting a problem, she/he raised his/her hand and was approached by a helper—either the experimenter or a research assistant. The helper answered the student's questions in a semi-standardized format. Specifically, the helper corrected the student's arithmetic mistakes (e.g., if the student multiplied 2 by 3 and got 5, the helper pointed out this error) or reminded him/her of the six possible transformations and how each was used. Helpers never gave strategic advice to students, such as suggesting which transformation to apply next or whether one method of solution was any better than another method. Students independently came up with their own choices for which transformations to apply. Also, students were encouraged to show all of their work by writing out their steps on all problems.

4.3. Materials

There were a maximum of 45 problems that students solved during the three problem-solving sessions. Students who did not finish a particular day's problems returned the following session and began where they had left off. As students worked at different paces, not all students had time to complete all 45 problems.

The session problems were carefully designed according to the following principles. Whenever possible, problems had integral coefficients and constants, to minimize any cognitive load issues and also to avoid biasing students, because of load issues, toward choosing one strategy over another. In addition and for similar reasons, all problems had integral solutions. The problems gradually increase in complexity, although more straightforward problems are presented in all three sessions to evaluate changes in students' approaches to them.

In addition, the problems used in this study were designed to give students maximal opportunities to demonstrate certain kinds of invention in their solution strategies. In particular, problems were created to give students the chance to invent new, potentially better solution procedures in three ways: **CHANGE IN VARIABLE**[5], **CANCEL TERMS**, and **DIVIDE BEFORE EXPANDING**. The **CHANGE IN VARIABLE** invention is the "change in variable" strategy that has been previously discussed. It results from an atypical use of the COMBINE transformation (e.g., adding $4(x + 1) + 2(x + 1)$ to get $6(x + 1)$) or the SUBTRACT FROM BOTH transformation (e.g., subtracting $2(x + 2)$ from both sides of an equation). The **CANCEL TERMS** invention, also discussed above, allows solvers to "cancel" terms as an initial solving step, and it results from an atypical use of the SUBTRACT FROM BOTH transformation. The **DIVIDE BEFORE EXPANDING** invention gives students the opportunity to divide an entire equation by a constant before expanding. Each of these atypical uses of transformations has the potential to result in a more efficient solution on some problems.[6]

Students took a post-test during the final session, which was identical to the pretest. Students were given 30 min to complete the post-test, and all students finished in the allotted time.

5. Analysis

All study instruments were graded by two independent coders, who subsequently met to resolve all disagreements. For each instrument, two scores were calculated: one for the percentage of problems completed correctly (e.g., those on which the students arrived at the correct numerical answer) and one for the percentage of problems completed without any errors in how transformations were applied, though arithmetic errors were allowed.

5.1. Measures

Students' solutions to post-test problems were coding for the presence of flexibility. Recall that flexibility was defined in terms of two subcomponents: knowledge of multiple solution procedures, and ability to invent. Variables were created for each of these capacities.

[5] Throughout this paper, inventions are indicated by bold and all capital text.

[6] A fourth invention involved multiplying the entire equation by 10 to "clear" decimals. Only one post-test problem gave students the opportunity to exhibit this invention, and no student came up with this strategy. As a result, this invention was not considered in further analysis.

Knowledge of multiple solution procedures was operationalized as follows. A subset of post-test problems (the last six problems) were identified as meeting the following criteria: Each problem could be solved in a number of ways, and some of these solution methods were at least as efficient, if not more efficient, than the standard algorithm. In other words, the final six post-test problems allowed students the opportunity to use multiple solution procedures, if students judged that deviating from the standard algorithm was a good idea. Looking only at these six post-test problems, the number that each student attempted that were solved with unique sequences of transformations were calculated. For example, if a student used the standard algorithm on all six problems, he would earn a "1" on this variable, indicating the presence of only one unique solution strategy on all six problems. If a student used a different solution method on each of the six problems, she would earn a "6" on this variable.

With respect to invention, a subset of 11 of the 19 post-test problems was identified as problems for which invention was possible. On some problems, students had the opportunity to use only one invention; for example, on problem 10, $4(x + 3) = 16x$, students were evaluated only as to whether they used the **DIVIDE BEFORE EXPANDING** invention. On other problems, students had the chance to use multiple inventions; for example, on problem 15, $3(x + 1) + 6(x + 1) + 6x + 9 = 6x + 9$, students could have used all three possible inventions (**DIVIDE BEFORE EXPANDING**, **CANCEL TERMS**, and **CHANGE IN VARIABLE**). Three variables were created to capture whether or not students invented—one for each of the three types of inventions. Each variable indicated the number of invention opportunities each student took advantage of, for each invention type. For example, students had the opportunity to use the **CHANGE IN VARIABLE** invention six times; a student who used this invention on only three of these problems would earn a score of 3 for this variable. In addition, an overall variable for invention was also created, indicating the total number of invention opportunities a student took advantage of on all 11 problems.

In addition to coding for flexibility, students' solutions were also coded for whether or not they discovered and used the standard solution method. For many problems, the standard method is the most efficient solution procedure; it was of interest to determine whether students discovered this method on their own. Knowledge of the standard solution method was assessed by looking for two features in students' solutions on the post-test. First, a solution was evaluated as to whether terms on each side of the equation were combined first, rather than being moved individually from one side to the other side. This feature is referred to here as "combine first," and it indicates (in the most typical case) whether students were able to efficiently transform an equation from its initial state to the form $ax + b = cx + b$. Students were able to demonstrate knowledge of this feature on nine post-test problems. The second feature, referred to as "move opposite", indicates whether students were able to efficiently transform an equation from the form $ax + b = cx + d$ to the solution state; students were able to demonstrate knowledge of this feature on six problems. Three variables were created for these features, indicating the number of eligible problems on which a student combined first, moved opposite, or both.

6. Results

Four students (three females and one male) were omitted from the analysis. One student withdrew from the study in the middle of the week and one student showed knowledge of formal equation solving techniques on the pre-test. In addition, two students who did not show knowledge of formal equation solving techniques on the pre-test nevertheless did show such knowledge in the first problem-solving session and were dropped.[7] Thus, of the 36 initially enrolled in this study, the work of 32 students—15 in the control group (7 male and 8 female) and 17 in the treatment group (8 male and 9 female)—was used in further analysis.

Unless otherwise indicated, treatment effects are explored with oneway ANOVAs, with pre-test score entered as a covariate. Eta squared (η^2) is used to report effect sizes, which can be interpreted as the amount of variance accounted for by the target variable.

The results are shown in Table 3 and discussed below.

6.1. Pre-test

Students did poorly on the pre-test ($M = 24\%$), indicating lack of prior knowledge of formal equation solving techniques. The few problems that students were able to complete correctly were the very straightforward ones from the beginning of the pre-test. On these initial problems, correct solutions were arrived at using various informal methods, including unwinding and guess-and-check. There were no significant treatment effects on the pre-test, indicating that random assignment of students to condition was done without bias.

6.2. Post-instruction test

Recall that students were given a short test immediately following instruction to assess learning of the six equation solving transformations. Students did very well on the post-instruction test ($M = 88\%$), indicating that the short period of instruction led to mastery of the equation solving transformations. There were also no significant treatment effects on the post-instruction test, indicating that instruction was given to both conditions without bias.

6.3. Problem-solving sessions

Students attempted an average of 36 problems during the three problem-solving sessions.[8] However, there were significant differences in how many problems each

[7] These two students revealed that they had received substantial tutoring from their parents in how to solve equations between the pre-test administration and the first problem-solving session.

[8] For the purposes of the analysis of students' work during the problem-solving sessions, treatment students' second attempts on session problems are considered as if they are responses to different problems. However, recall that treatment students' second attempts are not considered in either measure of flexibility.

Table 3
Results

Variable	All, $n = 32$	Treatment, $n = 17$	Control, $n = 15$	p-Value
First session assessments				
Pre-test (% correct)	24	26	22	
Post-instruction test (% correct)	88	85	90	
Session problems				
Problems attempted	36	31	42	***
% of problems with correct answers	83	85	81	
% of problems with no transformation errors	78	74	84	*
Post-test problems				
Problems attempted	19	19	19	
% of problems with correct answers	77	76	79	
% of problems with no transformation errors	91	90	92	
Flexibility: Knowledge of multiple procedures				
# of problems with unique strategies (out of 6)	2.8	3.2	2.4	*
Flexibility: Capacity for invention				
All Inventions (out of 19)	1.7	2.5	0.7	*
CHANGE IN VARIABLE (out of 6)	0.2	0.4	0.0	
CANCEL TERMS (out of 3)	0.2	0.1	0.3	
DIVIDE BEFORE EXPANDING (out of 10)	1.3	2.1	0.5	*
Knowledge of standard solution procedure				
Combine first (out of 9)	5.4	5.5	5.3	
Move opposite (out of 6)	1.7	1.2	2.3	
Both combine first and move opposite (out of 5)	2.1	2.0	2.2	

* $p < .05$.
*** $p < .001$.

group attempted. Over the three sessions, the control group solved approximately 11 more problems than did the treatment group: for the control group, $M = 42$, and for the treatment group, $M = 31$, $F(1,31) = 20.262$, $p < .001$, $\eta^2 = .34$. Since the conditions were otherwise identical, the most likely explanation for why the control group worked so much faster is that it took treatment students more time to complete a previously-solved problem using a different ordering of transformations, as the alternative ordering task requested, then it took the control group to solve an isomorphic problem.

Students in this study were quite successful on equations that were solved in the three problem-solving sessions. On average, students solved 78% of attempted problems without transformation errors and arrived at the correct solution on 83% of attempted problems.[9] The fact that such a high percentage of attempted problems were solved correctly is noteworthy, especially considering that students had no prior

[9] On occasion, students were able to arrive at the correct solution despite making transformation errors, particularly by guessing on some of the equations in the first problem-solving session. This accounts for why the percentage of session problems with correct answers is higher than the percentage with no transformation errors.

knowledge of formal solving procedures, did not receive any instruction in how to chain together transformations to solve equations, and did not see any worked-out examples.

While there were no significant treatment differences in the percentage of attempted problems solved correctly in the problem-solving sessions, there was a significant treatment difference in terms of the frequency of transformation errors in the problem-solving sessions. The control group made fewer mistakes than the treatment group: the treatment group solved only 74% of problems without transformation error, while the control group solved 84% without error, $F(1,31) = 4.278$, $p < .05$, $\eta^2 = .13$. Transformation use errors most likely emerged while treatment students attempted to use transformations in a different way to complete previously solved equations, as the alternative ordering tasks required.

6.4. Post-test

Students attempted an average of 18.7 of the 19 problems on the post-test. The rate in which students correctly used transformations on the post-test ($M = 91\%$) was somewhat higher than during the problem solving sessions; however, students were somewhat less likely to arrive at a correct answer on the post-test ($M = 77\%$). There were no significant treatment differences in the number of post-test problems attempted, the percentage of attempted problems solved without transformation error, or the percentage of problems solved with a correct numerical answer (see Table 3).

Note that this result indicates that even though the alternative ordering condition resulted in treatment students solving fewer problems (and with slightly more errors) on the booklet problems, on the post-test treatment students' performance (number of problems solved correctly and frequency of error) was identically to the control students.

6.5. Flexibility measures

6.5.1. Invention

Students used one of the three equation-solving inventions (**CHANGE IN VARIABLE**, **CANCEL TERMS**, or **DIVIDE BEFORE EXPANDING**) on an average of 9% of invention opportunities (approximately 1.7 of the 19 invention chances). While this may seem like a low rate for invention, it exceeds expectations as compared to the results of prior research using populations of much more experienced solvers. In particular, Lewis (1981) found that undergraduates were likely to invent in their solutions in only 9–12% of invention opportunities, with professional mathematicians inventing only slightly more frequently (approximately 20% of invention opportunities). It is reasonable to expect that sixth graders, as novices, would have lower rate of invention than undergraduates and a much lower rate than mathematicians. Considering the brief period of instruction (30 min) and practice (three 1-h sessions), the lack of prior knowledge, and the age of the study participants, the fact that solvers chose to invent new strategies on even 9% of invention opportunities is a surprising finding.

There were also significant treatment effects in how frequently students used inventions. Students in the treatment group were more likely to invent in their solutions on post-test problems, $F(1,31) = 5.152$, $p < .05$, $\eta^2 = .14$. Treatment students took advantage of 2.5 of the 19 invention opportunities, while control group students only took advantage of 0.7 invention opportunities. In terms of the three individual inventions, the treatment and control groups differed only in the **DIVIDE BEFORE EXPANDING** invention. Treatment students took advantage of 2.1 of the 10 invention opportunities (as compared to 0.5 invention opportunities in the control group); this difference was significant, $F(1,31) = 4.562$, $p < .05$, $\eta^2 = .13$.

6.5.2. Knowledge of multiple solution procedures

Recall that students' knowledge of multiple solution procedures was assessed by examining how often students used a unique sequence of transformations to solve the final six problems on the post-test. On average, students used the unique sequences of transformations to solve 2.8 of the six problems. Treatment students were somewhat more likely to complete these six problems using multiple solution procedures; on average, treatment students completed 3.2 of the six problems with unique solutions, as compared to 2.4 of the six problems in the control group, but this difference was not significant, $p = .146$. (The difference between treatment and control students on this measure was 0.56 standard deviation units, so the failure of this test to reach significance may be due to insufficient power.) To examine group differences on this measure in a different way, the continuous variable for knowledge of multiple solution procedures was recoded into a categorical variable, indicating whether students solved at least half of the final six post-test problems with unique sequences of transformations. Treatment students were significantly more likely to solve at least half of the six post-test problems with unique transformation sequences, $\chi^2(1, N = 32) = 4.394$, $p < .05$; 13 of the 17 treatment students relied mostly on unique step sequences, as compared to only 6 of 15 control group students.

6.6. Knowledge of standard solution procedure

Also of interest was whether students in this study discovered how to use the standard solution procedure. Recall that students were not provided any instruction on the standard solution procedure; all who discovered this algorithm did so on their own. Interestingly, most students did discover the standard solution procedure; 29 of the 32 participants used either the "combine first" or the "move opposite" strategies on at least one problem on the post-test.

Students on average used the "combine first" component of the standard solution procedure on 5.4 of nine possible problems and the "move opposite" strategy on 1.7 of six possible problems. On the five post-test problems where it was possible to use both the "combine first" *and* "move opposite" strategies, students did so on an average of 2.1 problems. There were no treatment differences on any of the measures assessing knowledge of the standard solution procedure. This indicates that treatment students were no less likely to make this discovery, despite the fact that the treatment group completed significantly fewer problems than the control group.

7. Discussion

In this study, students with little prior knowledge of equation solving learned, via minimal instruction and a few hours of practice, to be very successful solvers. In addition, some students discovered how to invent in their solutions, meaning that they showed the capacity to use solving transformations in atypical ways on some problems. The number of students who developed the capacity to invent was surprisingly large, considering performance of more mathematical experienced persons in prior research. Students' success and discovery of inventions is even more impressive considering the absence of worked-out examples during instruction.

More interestingly, differences emerged as a result of students' completion of alternative ordering tasks. The treatment and control groups did not differ in the accuracy of their solutions, both at the beginning of the study and at the conclusion of the study. However, those who were asked to solve previously completed equations using a different ordering of transformations were more likely to demonstrate the capacity to invent and to demonstrate knowledge of multiple solution procedures than those in the control group. These results suggest that the alternative ordering tasks were powerful in pushing treatment students to be more flexible solvers. Being asked to complete a problem again, using a different ordering of transformations, led (for many students) to the discovery of different, atypical ways in which transformations could be used, some of which turned out to be inventions. Treatment students' gains in flexibility came despite having completed significantly fewer problems than the control group during practice sessions. In addition, the gains in flexibility came without any cost to solution efficiency for the treatment group; despite having completed significantly fewer problems during the practice sessions, the treatment group was no less likely to discover the standard solution method.

This finding, that students can learn from a task such as the alternative ordering task, has not previously been reported in the literature on algebra learning. At the elementary school level, reform documents have long advocated the use of multiple and invented algorithms for solving arithmetic problems (National Council of Teachers of Mathematics, 1989, 2000). It has been suspected that allowing children to work with invented arithmetic algorithms, rather than being drilled in the use of standard algorithms, is beneficial toward their understanding of number (e.g., Carpenter et al., 1998; Carroll, 2000; e.g., Fuson et al., 1997). However, adapting these methods for post-arithmetic, symbolic mathematics has not been considered. The present study suggests that there are significant benefits to having students invent their own symbolic methods of solving equations and subsequently attempt to modify and refine these methods. As almost all students discovered the standard solution procedure on their own anyway, there was little cost in efficiency to allowing for discovery as opposed to explicitly teaching the most efficient solution strategy.

One possible critique of the present study is that its results are trivial; according to this viewpoint, the fact that students invented new strategies for problem solving is not surprising, given that the treatment specifically asked students to invent new strategies. In response to this critique, we note that the literature on mathematical problem solving offers ample reason to doubt that the alternative ordering task would have any effect at all.

First, this literature predicts that when a student discovers a solution method that enables her to solve a problem successfully, she will persist in using this solution method (e.g., Lewis, 1981, 1988; Schoenfeld, 1985). With continued practice, she will be even more likely to stick with her method that works, as using it will become more automatic and thus easier and faster to execute (Anderson, 1982; Anderson, 1992). Thus, participating in the alternative ordering treatment might be predicted to have no effect on students' strategy development; a treatment student may go through the motions and attempt to come up with another, different strategy, but she would only view this task as an exercise, not as a means toward learning a new approach. For such a student, there is no need to learn a new strategy; she already knows one that works and can be executed quickly and easily. In surprising contrast to this prediction, treatment students in the present study did learn from their second attempt; they modified their already-successful strategies based on what they observed from solving the same problem a second time.

Second, there is evidence from the problem solving literature that students, particularly novices, do not hold expert conceptions of what it means for a strategy to be better than another—that generating a more innovative solution method and recognizing it as better are quite distinct. Several studies have found that students hold idiosyncratic perceptions of what it means for one strategy to better than another, including subjective criteria such as personal like/dislike or the neatness and organization of the written strategy (Franke & Carey, 1997; McClain & Cobb, 2001; Star & Madnani, 2004). If a student's second attempt contained (what an expert would view as) innovative methods, it is not at all clear that she would recognize this and make use of these method in future attempts. (Furthermore, the task of determining which strategy was better was also complicated by the fact that students did not receive any feedback on their solution attempts in terms of right/wrong answers. So in addition to not being able to tell which solution method was better, students also may not have known whether one solution method was more likely to lead to the correct answer.)

In sum, the treatment students in this study were asked to solve an equation, and they typically generated a way that worked in their first attempt. Upon being asked to generate another strategy for the same problem, they generated a method that they perceived as different and that may have been better (although this may not have been apparent to the students). Nevertheless, on post-test measures, treatment students were more likely to use multiple solution strategies and also were more likely to invent in their solution. We find this to be quite striking, relative to what the literature might predict.

It is also important to emphasize that, given the absence of a delayed post-test, this study did not provide direct evidence that the treatment led to flexible problem solving in the long-term; only short-term gains were found. (Furthermore, there were some start-up costs associated with gains in short-term flexibility: Treatment students attempted fewer booklet problems and tended to make more errors during learning.) However, there is convincing evidence in the literature that short-term gains in flexibility have clear benefits for learning and performance in the long-term. For example, learners with knowledge of multiple strategies at pretest are more likely to learn from instructional interventions in both the short- and the long-term (Alibali, 1999; Siegler, 1995). More generally, the presence and benefits of multiple strategies has been linked to long-term gains in flexibility in a large range of domains, including locomotion (Adolph, Eppier, & Gibson, 1993), drawing (Karmiloff-Smith, 1990), serial recall (Siegler, 1989), spelling (Rittle-Johnson & Siegler, 1999), time-telling (Siegler & McGilly, 1989), and various kinds of elementary school mathematics problems (Alibali, 1999; Carpenter et al., 1998; Carroll, 2000; Fuson et al., 1997; Resnick, 1980; Resnick & Ford, 1981; Siegler, 1996; Siegler & Chen, 1998; Siegler & Jenkins, 1989; Siegler & Shrager, 1984). The present intervention reliably led to short-term gains in flexibility, and such gains have been linked in prior research to long-term benefits for learning.

7.1. Implications

The results of this study contribute to the growing literature on procedural flexibility in a number of ways. First, this work shifts emphasis from the description of flexibility, which has been the predominant focus in prior work, to the identification and evaluation of instructional techniques that may contribute to the development of flexibility. It has been found that experts are very flexible in their use of strategies (Dowker, 1992; Dowker, Flood, Griffiths, Harriss, & Hook, 1996); in addition, with increasing problem-solving experience, even young children become quite adaptive in their choices among strategies (Shrager & Siegler, 1998; Siegler, 1996; Siegler & Jenkins, 1989; Siegler & Shrager, 1984). Less is known about the kinds of tasks that may facilitate the development of flexibility. The results of this study, that the generation and comparison of multiple solution procedures can lead to increased flexibility, are a step in this direction. The present results are also consistent with recent work by Blöte et al. (2001), who found that a curriculum that focused on having students create and discuss their solution procedures and relate procedural problems to real-life contexts reliably led to increased procedural flexibility. Additional work is needed to identify and evaluate tasks and curricula that can foster the development of flexibility.

Second, this study broadens the domain in which flexibility is explored from arithmetic, which has been the almost exclusive focus in the field, to algebra. The processes by which flexibility develops in arithmetic and algebra may be quite different, for several reasons. Students' knowledge of arithmetic strategies is highly influenced by informal, out-of-school experiences, which is typically not the case with algebra. The procedures of algebra are longer and more intricate, meaning that the generation and comparison of multiple strategies imposes greater demands on cognitive resources. In addition, because older students typically are more capable of dealing with the abstraction inherent in algebra, it may be that the comparison and generation of multiple solution strategies is not as necessary in older students. The present results indicate that algebra learners *can* benefit from considering multiple solution procedures, in much the same way the elementary students do.

Furthermore, one concern about the use of tasks such as the alternative ordering task in algebra is that the discovery, comparison, and generation of alternative procedures will not necessarily lead students to discover the standard or modal way of solving problems. Standard algorithms are extremely powerful in algebra; students who fail to learn them may be at a disadvantage in later mathematics courses. This study found that almost all students at least partially discovered how to use the standard algorithm on their own, despite no direct instruction on this algorithm, minimal feedback, and only several hours of problem solving.

Acknowledgments

Thanks to Jack Smith, Bethany Rittle-Johnson, Anke Blöte, Deborah Ball, Magdalene Lampert, and several anonymous reviewers for their comments on earlier versions of this manuscript.

Appendix A. Equations attempted during problem-solving sessions

#	Treatment problem	Control problem
1.1	$3x = 3 + 12$	Same as treatment
1.2	$4x + 2 = 10$	Same as treatment
1.3	$4x = 2x + 12$	Same as treatment
1.4	$10x - 5x = 20$	Same as treatment
1.5	$4x + 10 = 2x + 16$	Same as treatment
1.6*	$4x + 10 = 2x + 16$	$6x + 9 = 3x + 12$
2.1	$6x + 8 = 2x + 16$	$15x + 10 = 5x + 20$
2.2*	Same as 2.1	Same as treatment
2.3	$3(x + 1) = 15$	$2(x + 3) = 10$
2.4*	Same as 2.3	Same as treatment
2.5*	Same as 2.3	$5(x + 2) = 20$
2.6	$4(x + 3) = 8x$	Same as treatment
2.7	$3(x + 2) + 9x = 3x + 15$	$2(x + 5) + 4x = 4 + 8x$
2.8*	Same as 2.7	Same as treatment
2.9*	Same as 2.7	$5(x + 3) + 10x = 35 + 5x$
2.10	$2(x + 4) + 6x = 8x + 4x$	Same as treatment
2.11	$4(x + 3) = 2(x + 3)$	$3(x + 2) = 6(x + 2)$
2.12*	Same as 2.11	Same as treatment
2.13	$9(x + 2) + 3(x + 2) + 3 = 6x + 9$	Same as treatment
3.1	$2(x + 3) + 4(x + 3) + 8x = 8x + 6$	Same as treatment
3.2	$6(x + 2) + 3(x + 2) + 6x = 9 + 6x$	Same as treatment
3.3	$2(x + 1) = 14$	Same as treatment
3.4*	Same as 3.3	$3(x + 2) = 21$
3.5*	Same as 3.3	$5(x + 3) = 25$
3.6	$2(x + 1) + 6(x + 1) = 4(x + 1)$	Same as treatment

(*continued on next page*)

Appendix A (*continued*)

#	Treatment problem	Control problem
3.7	$2(x+1)+6(x+1)+4(x+1) = 10x+8x$	Same as treatment
3.8*	Same as 3.7	$3(x+2)+6(x+2)+9(x+2) = 3x+6x$
3.9*	Same as 3.7	$5(x+1)+10(x+1)+5(x+1) = 5x+10x$
3.10	$2x+6=14$	Same as treatment
3.11	$2(x+1)+4(x+1)+6x+8 = 6(x+1)+8x+2$	Same as treatment
3.12*	Same as 3.11	$3(x+2)+6(x+2)+3+9x = 9(x+2)+6+6x$
3.13	$2(x+5)=4(x+5)$	Same as treatment
4.1	$2(x+3)+4x = 4x+6x+2$	Same as treatment
4.2*	Same as 4.1	$3(x+2)+6x+9x=18+6x$
4.3	$3(x+1)+6(x+1)+6x+9 = 6x+9$	Same as treatment
4.4	$0.1x+0.8=0.2x+0.3x$	Same as treatment
4.5*	Same as 4.4	$0.2x+0.4x=0.9+0.3x$
4.6*	Same as 4.4	$0.5x+0.3x=0.6x+0.8$
4.7	$4(x+2)+6x+10 = 2(x+2)+8(x+2) + 6x+4x+8$	$6(x+1)+9(x+1)+9+6x = 12(x+1)+9x+6x+6$
4.8*	Same as 4.7	Same as treatment
4.9*	Same as 4.7	$5(x+1)+15(x+1)+30+15x = 10(x+1)+15x+20+20x$
4.10	$2(x+3)+4x+8=4(x+2)+6x+2x$	Same as treatment
4.11	$0.5x+0.2=0.3(x+2)$	Same as treatment
4.12*	Same as 4.11	$0.3(x+5)+0.1x=0.3$
4.13	$0.3(x+1)+0.1(x+1) = 0.2(x+1)+0.4x+0.4$	Same as treatment

* Indicates problems on which treatment students engaged in the alternative ordering task.

References

Adolph, K. E., Eppier, M. A., & Gibson, E. J. (1993). Crawling versus walking: Infants perceptions of affordances for locomotion over sloping surfaces. *Child Development, 64*(4), 1158–1174.

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

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*(4), 369–406.

Anderson, J. R. (1992). Automaticity and the act* theory. *American Journal of Psychology, 105*(2), 165–180.

Anzai, Y., & Simon, H. A. (1979). The theory of learning by doing. *Psychological Review, 86*(2), 124–140.

Blöte, A. W., Klein, A. S., & Beishuizen, M. (2000). Mental computation and conceptual understanding. *Learning and Instruction, 10*, 221–247.

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

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*(1), 3–20.

Carroll, W. M. (2000). Invented computational procedures of students in a standards-based curriculum. *Journal of Mathematical Behavior, 18*(2), 111–121.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.). *Advanced in instructional psychology: Educational design and cognitive science* (Vol. 5, pp. 161–238). Mahwah, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H., & Bassok, M. (1989). Learning from examples via self-explanations. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 251–282). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

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

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

Dowker, A. (1992). Computational estimation strategies of professional mathematicians. *Journal for Research in Mathematics Education, 23*(1), 45–55.

Dowker, A., Flood, A., Griffiths, H., Harriss, L., & Hook, L. (1996). Estimation strategies of four groups. *Mathematical Cognition, 2*(2), 113–135.

Franke, M. L., & Carey, D. A. (1997). Young children's perceptions of mathematics in problem-solving environments. *Journal for Research in Mathematics Education, 28*(1), 8–25.

Fuson, K. C., Wearne, D., Hiebert, J. C., Murray, H. G., Human, P. G., Olivier, 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*(2), 130–162.

Gick, M. L. (1986). Problem-solving strategies. *Educational Psychologist, 21*(1–2), 99–120.

Hiebert, J., & Carpenter, T. (1992). Learning and teaching with understanding. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65–97). New York: Simon & Schuster Macmillan.

Hiebert, J., & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. *Cognition and Instruction, 14*(3), 251–283.

Karmiloff-Smith, A. (1990). Constraints on representational change: Evidence from children's drawing. *Cognition, 34*, 57–83.

Klein, A. S. (1998). Flexibilization of mental arithmetic strategies on a different knowledge base: The empty number line in a realistic versus gradual program design. Unpublished doctoral dissertation, Leiden University, Leiden, The Netherlands.

Krutetskii, V. A. (1976). *The psychology of mathematical abilities in school children (J. Teller, Trans.)*. Chicago: University of Chicago Press.

Lewis, C. (1981). Skill in algebra. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 85–110). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lewis, C. (1988). Why and how to learn why: Analysis-based generalization of procedures. *Cognitive Science, 12*, 211–256.

Luchins, A. S. (1942). Mechanization in problem solving. *Psychological Monographs, 54*(248).

Mawer, R., & Sweller, J. (1982). Effects of subgoal density and location on learning during problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(3), 252–259.

McClain, K., & Cobb, P. (2001). An analysis of development of sociomathematical norms in one first-grade classroom. *Journal for Research in Mathematics Education, 32*(3), 236–266.

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Pirolli, P., & Recker, M. (1994). Modeling individual differences in students' learning strategies. *The Journal of the Learning Sciences, 4*, 1–38.

Reimann, P., Schult, T. J., & Wichmann, S. (1993). Understanding and using worked-out examples: A computational model. In G. Strube & K. F. Wender (Eds.), *The cognitive psychology of knowledge*. Amsterdam: North-Holland.

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

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.

Resnick, L. B. (1980). The role of invention in the development of mathematical competence. In R. H. Kluwe & H. Spada (Eds.), *Developmental models of thinking* (pp. 213–244). New York: Academic Press.

Resnick, L. B., & Ford, W. W. (1981). *The psychology of mathematics for instruction*. Hillsdale, NJ: Lawrence Erlbaum.

Rittle-Johnson, B., & Siegler, R. (1999). Learning to spell: Variability, choice, and change in children's strategy use. *Child Development, 70*(2), 332–348.

Schoenfeld, A. (1985). *Mathematical problem solving*. Orlando: Academic Press.

Shrager, J., & Siegler, R. S. (1998). A model of children's strategy choices and strategy discoveries. *Psychological Science, 9*, 405–410.

Siegler, R. S. (1989). How children choose among serial recall strategies. *Child Development, 60*(1), 172–182.

Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology, 28*(3), 225–273.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Siegler, R. S., & Chen, Z. (1998). Development differences in rule learning: A microgenetic analysis. *Cognitive Psychology, 36*(3), 273–310.

Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Lawrence Erlbaum.

Siegler, R. S., & McGilly, K. (1989). Strategy choices in children's time-telling. In I. Levin & D. Zakay (Eds.). *Time and human cognition: A life-span perspective* (Vol. 59, pp. 185–218). Amsterdam: North-Holland.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Lawrence Erlbaum.

Simon, H. A., & Reed, S. K. (1976). Modeling strategy shifts in a problem-solving task. *Cognitive Psychology, 8*(1), 86–97.

Star, J. R. (2002). Re-conceptualizing procedural knowledge: Innovation and flexibility in equation solving (Doctoral dissertation, University of Michigan, 2001). *Dissertation Abstracts International, 62*, 3327A.

Star, J. R., & Madnani, J. (2004). Which way is best? Students' conceptions of optimal strategies for solving equations. In D. McDougall & J. Ross (Eds.), *Proceedings of the twenty-sixth annual meeting of the North American chapter of the International Group for the Psychology of Mathematics Education* (pp. 483–489). Toronto: University of Toronto.

Sweller, J. (1983). Control mechanisms in problem solving. *Memory and Cognition, 11*(1), 32–40.

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

Sweller, J., Mawer, R. F., & Ward, M. R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General, 4*, 639–661.

Zhu, X., & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction, 4*(3), 137–166.