DOES THE ELICITATION OF SELF-EXPLANATION OF CORRECT AND INCORRECT SOLUTIONS IMPROVE COLLEGE STUDENTS’ PROBLEM SOLVING ABILITIES?

By

LAURA A. CURRY

A THESIS PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE UNIVERSITY OF FLORIDA 2002
ACKNOWLEDGMENTS

My experience as a graduate student at the University of Florida has been an incredible opportunity for me to grow both personally and professionally. I attribute this fact to the many supportive and inspiring people I have met while here.

First, I would like to express my gratitude to the faculty, staff, and my peers in the developmental psychology department, for creating a supportive and stimulating academic environment from which I been able to succeed.

I am extremely thankful to my advisor, Dr. Jennifer Woolard, for her unwavering support and guidance throughout the past four years. I am especially indebted to her for believing in me through times of personal doubt and adversity. Jen has been a tremendous role model for me. She is a talented and gifted academic and a caring and gracious individual. I hope to have the opportunity to work with her again sometime in the future.

Special thanks also go to my committee members. Dr. Shari Ellis offered her expertise and insight, especially
with regard to the design and development of this study. With the help and encouragement of Dr. Scott Miller, I have learned the finer points of APA style. For the first time in my life, I can be proud of my writing. I have also benefited greatly under the tutelage of Dr. James Algina and am continually grateful for his expertise in statistics and his infinite patience with all of my questions.

I would also like to offer thanks to my fellow graduate students in the developmental psychology department for continually offering support, friendship, and diversions. I offer special thanks to Joann Benigno and Angi Semegon for providing me with shoulders and hugs, giving me hope during tough times, and celebrating with me during the good times.

I am eternally grateful to Henry Morgenstern, for changing my life in ways I had never imagined or considered. He has inspired me and has given me the courage to be my best.

Finally, words can not express the heartfelt gratitude I offer my parents. It is because of their continued unconditional support, love, and encouragement that I have succeeded in this endeavor. My parents believed in me when I could not believe in myself, and provided me with strength and perspective when my own failed. Because of
them I am not afraid to fail. I know that there are worse things in life. And I look forward to the future with hope.
TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................... ii
LIST OF FIGURES ........................................ vii
ABSTRACT ............................................... viii

CHAPTER

1 INTRODUCTION AND BACKGROUND ......................... 1
Cognitive Models/Strategies ................................ 2
Conceptual versus Procedural Knowledge ................... 3
Differences in Classroom Instruction between American and Asian Schools ................. 5
Effects of Self-Explanation ................................ 6
Correct versus Correct and Incorrect Strategies .......... 11
Effects of Feedback ........................................ 12
Current Study ............................................. 14
Hypotheses ............................................... 15
The Selected Task ........................................ 16

2 METHOD .................................................. 18
Participants ............................................. 18
Measures ................................................ 19
Procedure .............................................. 22
Design .................................................. 25

3 RESULTS .................................................. 30
Covariates ................................................. 32
Overall Practice Effects .................................. 33
Group Effects ............................................ 36
Feedback Effects ......................................... 40
Self-Explanation Effects .................................. 42
Combined Effects of Feedback and Self-Explanation .... 43
Summary ................................................. 44
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Total posttest mean scores by experimental group</td>
<td>37</td>
</tr>
<tr>
<td>2</td>
<td>Posttest mean scores by Group and Problem Type</td>
<td>39</td>
</tr>
<tr>
<td>3</td>
<td>Posttest mean scores by Feedback Group and Problem Type</td>
<td>41</td>
</tr>
</tbody>
</table>
The effects of giving feedback and eliciting self-explanation on algebra word-problem solving were studied. Eighty college students were randomly assigned to groups with different feedback (no feedback, ambiguous feedback, feedback) and self-explanation (correct, correct and incorrect) conditions to test for effects of practice, group condition, and problem type. Students given feedback and those prompted to explain both correct and incorrect solutions were expected to outperform others. Self-explanation effects were expected to be most apparent for problems that required greater conceptual understanding; however this was not observed. Significant effects were found for practice, feedback, problem type, and for self-
explanation within the simplest problem type. Findings offer insights that could prove valuable to educators in selecting task-appropriate instructional strategies.
CHAPTER 1
INTRODUCTION AND BACKGROUND

There are various instructional methods that can be used with at least some level of success to facilitate learning and understanding. These methods include, but are not limited to, self-explanation (Chi, 2000), collaborative problem solving (Ellis, Klahr, & Siegler, 1993), scaffolding (Vygotsky, 1978), reciprocal teaching (Brown & Palinscar, 1989), and learning from worked out examples (Mwangi & Sweller, 1998). Developmental and educational psychologists are particularly interested in the cognitive processes affected by these methods, the varying effectiveness of each across different domains and ages of targeted students, and the mechanisms that are associated with the learning that results from the utilization of each. Although these instructional techniques are different in form, they are similar in an important way. Each one encourages the student to engage in a learning session during which knowledge is actively processed, and mental models and schema are constructed and reconstructed.

The effects of feedback and self-explanation have been examined under various conditions, and within various
domains (Alibali, 1999; Chi, de Leeuw, Chiu, & LaVancher, 1994; Ellis, 1997; Mwangi & Sweller, 1998; Neuman & Schwarz, 1998; Tudge, Winterhoff, & Hogan, 1996). Because both have shown to have advantageous effects under many circumstances, they were used together in this study of algebra problem solving. To extend prior research, the self-explanation of correct as well as incorrect solutions was elicited and compared to the condition in which only the correct answer was self-explained. It was expected that students who received feedback and were asked to explain both correct and incorrect solutions would demonstrate the most improvement in solving algebra word problems.

**Cognitive Models/Strategies**

In order to solve a novel problem, one must have a repertoire of strategies available from which to choose. How are these strategies developed? Various models and processes have been proposed to explain how strategies are generated and selected in the presence of a given problem. These models have often been categorized as either metacognitive or associative (Crowley, Shrager, & Siegler, 1997). Metacognitive models present strategy generation and selection as explicit processes (Anderson, 1996), whereas associative models depict these processes as being almost entirely automatic and implicit (Karmiloff-Smith, 1992;
Siegler & Shrager, 1984). Both types of models have merits, however Crowley and colleagues propose that neither quite captures the reality of developmental changes. As a result, Crowley et al. propose a model that reflects a “competitive negotiation” between metacognitive and associative representations of knowledge related to a given problem. That is, associative or implicit strategy choice processes coexist with and are supplemented by metacognitive or explicit processes. Consistent with this model, one could expect a student to use associative strengths when selecting a strategy unless the metacognitive system somehow perceives additional data and becomes involved. This additional data can be gained from outside feedback or self-monitoring processes such as self-explanation.

**Conceptual versus Procedural Knowledge**

The strategies one has available to solve a given problem are limited by one’s conceptual and procedural knowledge. According to Rittle-Johnson and Alibali (1999), conceptual knowledge is defined as “explicit or implicit understanding of the principles that govern a domain and of the interrelations between pieces of knowledge in a domain,” and procedural knowledge may be defined as “action sequences for solving problems” (p. 175). This distinction is especially relevant in the area of algebra. For example,
a student could be proficient at manipulating various given algebraic equations, thereby demonstrating a high level of procedural knowledge. The same student, however, may be unable to translate a given word problem into an appropriate equation, which would indicate a lack of conceptual understanding of the problem.

Any given instructional technique might affect a student’s conceptual understanding and procedural skills differently. As mentioned previously, evidence suggests that the use of self-explanation is beneficial to learning. However, some researchers have proposed that it is less effective in teaching procedural tasks than in stimulating conceptual understanding (Mwangi & Sweller, 1998; Nathan, Mertz, & Ryan, 1994). Mwangi and Sweller suggest two reasons for this. First, although a student may be able to perform a series of mathematical manipulations, the verbal expression of this task requires different skills than mathematical expression, and translation may be difficult. Second, if we consider the fact that tasks that require the student to perform a set of mathematical manipulations require a high level of cognitive resources, then the additional demand on these resources associated with self-explanation may detract from rather than facilitate learning. Sweller (1989) has also found that the level of
cognitive load is an important factor to consider when determining the effectiveness of learning from worked examples and instructional techniques that require the student to split his attention between sources of information in order to integrate the presented material.

**Differences in Classroom Instruction between American and Asian Schools**

Students gain conceptual understanding by examining incorrect (by drawing attention to knowledge gaps and discrepancies) as well as correct problem-solving strategies (Nathan et al., 1994). However, the process of education in the United States focuses on correct strategies and algorithms, while almost completely ignoring the existence of those that are incorrect. Whether learning occurs in a classroom environment (Stevenson & Stigler, 1992) or with a tutor in a one-on-one environment (Chi, 1996), U.S. teachers generally lead the student through the correct algorithms to solve problems. When a student mistakenly goes down an alternative, incorrect path, the teacher redirects that student, rather than allowing him or her to follow through with the erroneous strategy. The presumption behind this tactic is that by reinforcing the correct strategy it will become more salient to the student. In contrast, educators in Japan will often present
a new problem to students and allow them to develop strategies while working in small groups (Hatano & Inagaki, 1991, 1998; Stigler & Fernandez, 1995). Then, each strategy is examined and as a class students discuss the merits of each. The rationale here is that by analyzing several potential strategies, including incorrect ones, students are able to correct misconceptions and gain a deeper understanding of the underlying concepts imbedded in each problem. As Stevenson and Stigler explain the effectiveness of Chinese and Japanese instructional techniques, “Discussing errors helps to clarify misunderstandings, encourage argument and justification, and involve students in the exciting quest of assessing the strengths and weaknesses of the various alternative solutions that have been proposed” (1992, p. 191). Clearly, these techniques encourage the students to engage in active and constructive processing of knowledge.

**Effects of Self-Explanation**

Although the results of several studies that examined the effect of self-explanation on learning processes are varied, we can confidently posit that there are at least some circumstances under which this method proves beneficial. Chi and others have found that the spontaneous use of self-explanation (Chi, Bassok, Lewis, Reimann, &
Glaser, 1989; Ferguson-Hessler & de Jong, 1990) and also the elicitation of self-explanation (Chi et al., 1994) are associated with enhanced learning in the domains of physics and biology, respectively. Interestingly, however, Nathan et. al., (1994) found that elicited self-explanation resulted in greater test improvement for algebraic tasks in which conceptual reasoning was necessary, but not for problems requiring only procedural equation manipulation. There is also evidence that the elicitation of self-explanation enhances learning from worked-out examples (Renkl, Stark, Gruber, & Mandl, 1998), although Neuman and Schwarz (1998) suggest that only self-explanations that provide “deep structural explanations,” (p. 20) such as inference or clarification lead to improvements in analogical problem solving. In contrast, Mwangi and Sweller (1998) conducted a study with third graders, and found that the elicitation of self-explanation had no significant effect on learning to solve “compare word” problems. Thus, the use of self-explanation is not necessarily equally effective for all students under all circumstances.

Chi and colleagues (Chi, 1996; Chi et. al., 1989; Chi et. al., 1994; Chi & Van Lehn, 1991) have established various domains and conditions under which self-explanation facilitates learning. In the earliest of their studies
(1989), it was noted that university students who spontaneously generated self-explanations while solving physics problems outperformed other students. Subsequently, a study was conducted in which self-explanation condition was randomly assigned such that self-explanations were elicited from students in the experimental group (Chi et al., 1994). In this case, the participants were eighth graders, and the task was to read and answer questions related to an expository biology text. Although no main effects were found, students who self-explained outperformed others when answering more difficult questions. The authors concluded that self-explanations help to generate inferences regarding information not explicitly stated in the text.

It seems that inference generation is a key component of self-explanations; however Chi (1996) proposes that it is not the only, and perhaps not the most important, resulting benefit. Upon closer examination of the self-explanations of individual participants of the biology text study (Chi et al., 1994), Chi observed that these explanations seemed to foster the refinement and repair of mental models. In this case, it seems that self-explanation helped students to gain a deeper conceptual understanding of the subject matter.
Nathan and others (1994) also studied the effects of elicited self-explanation on the algebraic problem solving skills of university students. In this study, the tasks were to solve for an unknown variable (procedural manipulation) and to express algebraically a provided story problem (conceptual processing). Cognitive load condition (high or low) was also manipulated. Results revealed that self-explanation was beneficial only for the condition in which the task was conceptual based and cognitive load was low.

It is also not clear whether the advantages of self-explanation are consistent throughout the lifespan. Most related studies to date have involved older children or university students. Sweller (1989) maintains that learning is facilitated by the reduction of cognitive load. Self-explanations actually add to the demands on cognitive resources, during problem solving or studying. We also know that cognitive capacity increases with age throughout childhood. Therefore, it seems possible that advantages of this elicitation increase as scientific reasoning skills and cognitive capacity develop. This possibility would explain why Mwangi and Sweller (1998) found no significant effect of self-explanation in third graders. At that age, perhaps the additional cognitive demands of self-
explanation negate any potential advantages of this process.

Siegler (1995) also examined the effects of eliciting explanations from younger children while solving number conservation problems. Forty-five 5-year-olds were presented with problems. Children in the first group were given feedback as to whether their answers were correct or incorrect ("feedback only"). Children in the second group were asked to explain their own reasoning, then they received feedback ("feedback plus explain-own-reasoning"). And, the last group of children first received feedback, and then were asked to explain the experimenter’s reasoning ("feedback plus explain-experimenter’s-reasoning"). Siegler found that children in the last group outperformed the other children. He concluded that it was not the act of explaining per se, but rather the act of explaining another’s more advanced reasoning that facilitated learning in this case. While this is certainly a possible explanation for these results, it is also possible that the difference in the order in which feedback was received influenced the effects of eliciting explanations. Perhaps the results would have been different if children in the second group first received feedback and then were asked to explain why they thought the correct answer was correct.
Correct versus Correct and Incorrect Strategies

As discussed, research has established that certain types of self-explanation can lead to improved performance of some tasks. Both inference generation and conceptual understanding have been affected, although not uniformly over all ages and tasks. However, it still is not clear how or why self-explanation is effective. For example, we do not yet fully understand the effects that this elicitation has on the set of potential strategies available to a student during the problem-solving process. Perhaps self-explanation emphasizes faulty logic within strategies, perhaps it simply emphasizes cognitive conflict, or simply helps to identify gaps in knowledge.

The effects of self-explanation as they pertain to the explanation of correct (or at least what the student deems to be correct) strategies or interpretations have been examined. Whereas the use of self-explanation in this manner may lead to the generation of new strategies it may not be completely effective in the removal of existing incorrect strategies. I propose that the additional elicitation of explanation as to why a given strategy or interpretation is incorrect may further enhance learning by accentuating the fallacious basis of such a solution. Therefore, self-explanations of incorrect strategies may
lead to removal of or adjustment to those incorrect strategies.

It follows, then, that students who are asked to explain both the merits of correct solutions and the deficits of incorrect solutions would gain additional conceptual understanding. As is indicated by the results of research conducted by Nathan et al. (1994), it is possible that the elicitation of self-explanation only improves the learning of that which requires conceptual knowledge. Conceptual knowledge is often measured by evaluating verbal explanations of solutions (Rittle-Johnson & Siegler, 1998). Conceptual knowledge can also be assessed by performance on tasks that require a high level of conceptual understanding, as was done in Nathan and others’ study. Therefore, we can expect that advantages due to the elicitation of both correct and incorrect solutions will be evident when analyzing changes in verbalized strategies more so than when analyzing procedural performance.

Effects of Feedback

According to Piaget (1972) an important mechanism of cognitive development is equilibration. That is, when an individual is confronted with cognitive conflict between existing knowledge and reality that seems to contradict that knowledge, he or she is motivated to modify that
knowledge in order to maintain cognitive equilibrium. However, while this conflict may be necessary for development to occur, is it sufficient?

Research has supported the belief that feedback also plays a significant role in cognitive development. Tudge and others (1996) examined the impact of feedback on collaborative problem solving and found that the performance of 6- to 9-year-olds who received feedback during the problem solving process improved significantly more than children who did not. In this case, strategy development was examined while students tried to predict results of balance beam manipulations. Further, they discovered that when children received feedback, the presence of a partner actually hindered problem solving (i.e., in some cases children adopted incorrect strategies from their partners.)

Ellis, Siegler, and Klahr (unpublished) conducted research in which feedback led to greater accuracy among fifth graders when comparing decimal numbers (as cited in Ellis, 1995). However, unlike the results of Tudge et al. (1996), collaboration also improved performance in both feedback and no feedback conditions. Ellis et al. found that while some of the children in the no feedback condition generated new incorrect strategies, only those in
the feedback condition were able to generate new correct strategies. Further, with regard to children in the feedback condition, those working with partners were roughly twice as likely to generate new correct strategies as those working with partners were roughly twice as likely to generate new correct strategies as those working alone.

Lastly, Alibali (1999) conducted a study in which she examined more closely the process by which third- and fourth-grade children change strategies. In looking at both strategies explicitly verbalized and those observable through gesture, she found that while feedback did not affect overall strategy generation it did affect generated verbal strategies. That is, she concluded that the effect of feedback was to motivate the children to verbalize their newly generated strategies.

**Current Study**

This study was conducted in order to examine the ways in which feedback and self-explanation condition affects students’ performance on algebra problem-solving tasks. First, the individual effects of feedback on performance were analyzed and compared to results of prior studies. It was expected that students who received feedback would show greater improvement between pre- and posttests than students who did not receive feedback. Next, the combined
effects of feedback and self-explanation of correct solutions, versus feedback with self-explanation of both correct and incorrect (or alternative) solutions, were analyzed. My preliminary hypothesis was that while self-explanation of correct solutions would likely increase the salience of correct algorithms or strategies, it would not address underlying misconceptions that the student may have. Conversely, however, I suspected that by also encouraging the student to self-explain why alternative solutions are inadequate, one would be more likely to address these misconceptions. Increased conceptual understanding would be attained, and inadequate solutions would be removed from the set of potential strategies held by the student for that problem type. Therefore, it was expected that students who received feedback would make greater improvements in performance than those who did not, and I predicted that students who received feedback and explained both correct and incorrect solutions would make the furthest gains in conceptual understanding.

**Hypotheses**

H1. Practice session will lead to improved posttest performance (compared to pretest performance), in general.
H2. Students in the experimental groups will outperform students in the control group.
H3. Feedback will positively affect posttest performance.

H4. Self-explanation of correct and incorrect solutions will positively affect posttest performance.

H5. There will be group by problem type interaction effects. Specifically, the self-explanation experimental condition will have significant positive effects on problems requiring conceptual knowledge (i.e., simple-indirect problem types). Experimental conditions will have no significant effect on simple-direct nor complex problems.

H6. Combined experimental positive significant effects will be significant. That is, students who receive feedback and are instructed to explain both correct and incorrect solutions will show significant improvement in test scores from pretest to posttest.

The Selected Task

It is widely recognized that students of all ages generally have difficulty solving algebra word problems. More specifically, a common error has been detected with regard to word problems that require the student to write an expression that represents a comparison of quantities between two variables. This error is commonly referred to as the “variable-reversal error.” For example, “There are six times as many students as there are professors at the
university," is a compare problem that can be algebraically represented by the equation \( S = 6P \). However, students often make the mistake of expressing the comparison as \( P = 6S \). Since this error has been well established in past research (Bernardo & Okagaki, 1994; Clement, 1982), I have chosen to use this type of problem in my study.
Participants

Participants included 80 students enrolled in an introductory psychology course at the University of Florida (60 females and 20 males). Students were recruited via Experimetrix experiment sign-up software and received credit toward a course requirement in return for their participation. Students were randomly assigned to one of four group conditions, resulting in 20 participants in each. However, SAT scores were missing for 3 participants. Therefore, for analyses in which SAT scores were used as covariates, $n = 18$ and $n = 19$ for the “explain correct and incorrect” and “ambiguous feedback” groups, respectively (group descriptions are included in the procedure subsection). Participant age ranged from 18.43 to 36.18 years old ($M = 19.73, SD = 2.05$) with most students between 18 and 20 years old. The ethnic breakdown of the group was 50 Caucasian, 12 African American, 10 Hispanic, 6 Asian, and 2 Other. All participants were treated in accordance with the ethical standards of APA.
Measures

Background Information

Participants were asked to complete a questionnaire requesting background information including gender, date of birth, and verbal and math Scholastic Achievement Test (SAT) scores.

Algebra pretest and posttest

These measures consisted of 14 multiple-choice algebra word problems, and were used to assess algebra problem-solving abilities. For each problem a written expression describing a comparison between two variables was presented. Each of four multiple-choice answers was presented in the form of an algebraic equation. Under each answer, a space was provided so that the participant could indicate how certain he or she was that the answer selected was the correct one (from 0 to 100%). Both tests consisted of four simple-direct, five simple-indirect, and five complex problems. All participants were given identical tests and both the pre- and posttests consisted of problem types presented in the same order. The intent was that the tests would be identical in form, and only different in surface features. For all problem types, the answer was a ratio between two variables. An example of each problem type and brief explanation is offered below:
In the case of the simple-direct problems, one variable may be expressed directly by multiplying the given factor (explicitly stated in the problem) by the other variable.

Example: On a nearby farm, the number of goats is five times the number of cows.

Simple-indirect problems are those that require an additional mathematical operation before a factor can be applied to the resulting equation.

Example: The student tickets are 40% less than the price of the general admission tickets. (Here, the student must subtract 40% from 100%, then apply the result in decimal form, or .60, to the variable representing the general admission tickets.)

And, complex problems are those in which two factors are stated in the problem. These factors must be appropriately applied to the two given variables.

Example: There are three boys for every two girls in the class.

Because correct equations for the simple-direct and complex problems can be derived directly from the information given in the associated problem, and because once an equation format is provided for one problem, it can easily be transferred to analogous problems (simply by
substituting the new variables for the old ones), these problems are categorized as requiring procedural knowledge. The simple-indirect problems are not as straightforward, however. These problems require additional manipulation of the information provided, and are therefore categorized as requiring more conceptual understanding. For example, in the complex problem above, the additional step is to subtract 40% from 100% and translate that result to decimal form, or .60. If, however, the problem had stated that the student tickets were 40% more than the general admission tickets, then the student would have to know to add 40% to 100%, or to apply a factor of 1.4 to the general admission variable.

**Algebra directed practice**

This measure consisted of 10 algebra comparison problems very much like those in the pre- and posttests. Three problems were of the simple-direct type, four were of the simple-indirect type, and three were of the complex type problems. Problems were presented in the same order for all participants. During the practice session multiple-choice answers were not provided. Instead, students were asked to provide algebraic equations on their own. This measure served two functions. First, it was during these sessions that the experimental conditions were applied.
Second, strategies used by students were recorded and levels of understanding were assessed during these sessions. These data will be used in future analyses.

**Procedure**

Each participant was tested individually by a trained experimenter in a laboratory in Walker Hall. After completing a background questionnaire, participants completed an algebra pretest. Specifically, they were asked to select from a set of four alternatives the algebraic equation that they thought most accurately represented the corresponding expression written in words. Further, they were instructed to indicate the certainty (0 to 100%) with which they believed that the equation was correct. There was no time limit imposed on the completion of this measure.

Following the pretest, students took part in a directed practice session (individually), during which all student work was written on a chalkboard and was videotaped. The experimenter introduced algebra directed practice session problems as follows:

*You do not have to solve the problems that follow. Simply write an expression that best represents each comparison as it is written. For each problem, you will be asked to explain the strategy that you used in developing each expression, and will be asked to explain why this expression does (or does not) represent the written comparison.*
Participants were randomly assigned to one of four conditions. Those assigned to the first condition ("Control") did not receive feedback with regard to the accuracy of each of their expressions, and were asked to explain why the expression that they provided for each problem was correct. In contrast, participants assigned to condition two ("Ambiguous Feedback") derived their own equations, but were also given an alternative solution by the experimenter. If the student’s answer was correct the experimenter provided an incorrect solution whereas if the student’s answer was incorrect, the experimenter provided a correct solution. After both equations were presented, the participant was told that one solution was correct and one was incorrect (the "ambiguous" feedback). They were not told which was which. Then, they were asked to explain why the answer they believed to be incorrect did not accurately represent the expression written in words. They were also asked to explain why the other equation did accurately represent the expression written in words.

Participants assigned to condition three ("Explain Correct") received feedback with regard to the accuracy of their derived equations. If a student’s equation was correct he or she was asked to explain why it was correct. If the equation was incorrect, the experimenter provided a
correct expression and the student was asked to explain why he or she thought that the provided equation was correct. Participants assigned to the final condition ("Explain Correct & Incorrect") also received feedback, but were given an alternative equation regardless of the correctness of their own equations. The alternative equation provided was accurate if the student’s equation was incorrect and was inaccurate if the student’s equation was correct. Further, these participants were asked first to explain why the incorrect equation was in fact incorrect and then were asked to explain why the correct equation was correct.

The following summary of group descriptions may be used for quick reference:

Control: No feedback was given; students were asked to explain why they thought their solution was correct.

Ambiguous Feedback: The experimenter provided the students with an alternative solution; students were told that one was correct and one was incorrect; they were asked to select the one they believed to be incorrect and explain why, then to explain why they believed the other to be correct.

Explain Correct: Explicit feedback was given; correct solutions were provided by the experimenter, when
necessary; students were asked to explain why the correct solution was correct.

Explain Correct & Incorrect: An alternative solution was provided by the experimenter; explicit feedback was provided; students were asked to explain why the incorrect solution was incorrect, then why the correct solution was correct.

After the directed practice session participants were asked to take an algebra posttest that was identical in form to the pretest, but contained different problems. Again, no time limit was imposed.

**Design**

Originally, we intended to analyze the data as if the experimental procedure reflected the manipulation of two between subjects variables, feedback (whether the participant received it or not) and self-explanation condition (explain correct or one’s own versus explain correct and incorrect, or one’s own and an alternative solution). However, upon further consideration, we realized that the four experimental conditions could not be simplified in that manner. Specifically, the group two condition was initially classified as one in which participants received no feedback. To more accurately
describe this condition, it was reclassified as one in which “ambiguous” feedback is provided.

Several types of analyses were used to evaluate the hypotheses. The first type includes univariate analyses of variance (ANOVAs) and analyses of covariance (ANCOVAs) in which total test scores were used. In addition to these, multivariate analyses of variance (MANOVAs) and covariance (MANCOVAs) with test scores by problem type were performed. The reason for this strategy is that the tests consisted of 14 problems: four simple-direct, five simple-indirect, and five complex. Total scores were out of 14; however, scores by problem type were recorded as percent correct out of 100 percent. Since the total mean of scores by problem type is not necessarily equivalent to total score, it was preferable to use total scores when analyzing data across all problem types. Mixed model MANCOVAs were used to examine interaction effects between group and problem type. In these cases group was the between subjects variable, and problem type was the within subjects variable. Following is a brief explanation of the analysis used for each stated hypothesis:

H1: The directed practice session will lead to improved posttest performance (compared to pretest performance), in general. To determine whether the practice
session had any general effect on posttest performance, a repeated measure (time - pre/post) ANOVA was performed. Next, a repeated measures 2 (time) by 3 (Problem Type) MANCOVA with pretest scores by problem type as covariates was used to examine problem type, and time by problem type interaction effects. Then, practice session effects within each group condition were analyzed by performing four individual (one for each group) repeated measure (time) ANOVAs, based on total test scores. Finally, practice effects were analyzed within each group and problem type, with the use of 12 individual repeated measure (time) ANOVAs.

H2: Students in experimental groups will outperform students in the control group. Group condition effect on total posttest performance, was examined with a one-way (Group) ANCOVA with total pre-test score and SAT math and verbal scores as covariates.

H3: Feedback will positively affect posttest performance. In order to understand the effect that feedback has on problem solving, total posttest mean scores for students in the Control group were compared to scores for students in the Explain Correct group (received feedback,) with the use of a one-way (Feedback) ANCOVA with total pretest and SAT scores as the covariates, and total
posttest score as the dependent variable. A mixed model 3 (Problem Type) x 2 (Group) MANCOVA with pretest scores by problem type as the covariates was also performed to investigate any group by problem type interaction effects.

H4: Self-explanation of correct and incorrect solutions will positively affect posttest performance, in general. Mean scores for the Explain Correct and Explain Correct and Incorrect groups were compared to test this hypothesis. A one-way (Self-explanation Group) ANCOVA with total pretest and SAT scores as the covariates, and a mixed model 3 (Problem Type) x 2 (SE Group) MANCOVA with pretest scores by problem type were used to examine the effects of self-explanation condition.

H5: Students who self-explained both correct and incorrect solutions will demonstrate greater gains in performance from pretest to posttest on problems that require conceptual knowledge, than students who explain correct (or “own”) solutions only. (i.e., The self-explanation experimental condition would have significant positive effects on simple-indirect problems.) Self-explanation condition effects were examined by performing an individual one-way (Group) ANCOVA with simple-indirect problem type pretest score as covariate and simple-indirect posttest score as the dependent variable.
H6: Combined experimental conditions would lead to significant improvements in performance. First, a one-way (Group) ANCOVA with total pretest and SAT scores as covariates, was performed to determine combined effects on total posttest performance. The posttest means for the Control group and Explain Correct and Incorrect group were compared. A mixed model 3 (Problem Type) by 2 (Group) MANCOVA with pretest scores by problem type as the covariates was used to compare combined experimental condition (Feedback and Self-explanation) by group interaction effects. Three individual one-way (Group) ANCOVAs were performed for each problem type with respective pretest scores as covariates to investigate simple main effects of experimental condition (Group).
The data were analyzed several different ways to determine the effects of the various group conditions. Both total test scores for pre- and posttests and scores by problem type (simple-direct, simple-indirect, and complex) were examined. The total scores represented the number correct from 0 to 14. However, since there were an unequal number of problems by problem type (4 simple-direct, 5 simple-indirect, and 5 complex), scores by problem type were based on percent correct from 0 to 100 percent.

The means and standard deviations of test scores by group and problem type are presented in Table 1. (Students’ total scores ranged from 1 to 13 on the pretest and from 0 to 14 on the posttest.) Note that pretest means by group and problem type are noticeably unequal. In fact, a one-way (Group) ANOVA revealed significant group differences in pretest scores. This must be considered when comparing posttest scores by group. Consequently, reported posttest means have been adjusted for random differences in covariate mean values, when ANCOVA results are presented. Total posttest scores reveal improved performance for all
groups, with greater improvement for students who received feedback. Examination of subscores, however, indicates possible interaction effects between group and problem type.

Table 1  Test Scores by Group

<table>
<thead>
<tr>
<th>Test by Problem Type</th>
<th>Group 1a</th>
<th>Group 2b</th>
<th>Group 3c</th>
<th>Group 4d</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple-direct</td>
<td>78.8 (18.6)</td>
<td>78.8 (26.0)</td>
<td>73.8 (20.6)</td>
<td>75.0 (28.1)</td>
<td>76.6 (23.3)</td>
</tr>
<tr>
<td>Simple-indirect</td>
<td>49.0 (35.8)</td>
<td>38.0 (31.1)</td>
<td>44.0 (30.9)</td>
<td>42.0 (33.3)</td>
<td>27.0 (27.7)</td>
</tr>
<tr>
<td>Complex</td>
<td>27.0 (27.7)</td>
<td>34.0 (30.5)</td>
<td>20.0 (19.5)</td>
<td>44.0 (41.9)</td>
<td>31.3 (31.6)</td>
</tr>
<tr>
<td>Total pretest</td>
<td>7.0 (2.9)</td>
<td>6.7 (2.6)</td>
<td>5.8 (2.8)</td>
<td>7.4 (4.0)</td>
<td>6.7 (3.2)</td>
</tr>
<tr>
<td>Posttests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple-direct</td>
<td>81.3 (26.8)</td>
<td>72.5 (31.3)</td>
<td>63.8 (31.9)</td>
<td>86.3 (25.0)</td>
<td>76.0 (39.6)</td>
</tr>
<tr>
<td>Simple-indirect</td>
<td>53.0 (25.4)</td>
<td>45.0 (30.4)</td>
<td>49.0 (27.9)</td>
<td>54.0 (26.1)</td>
<td>50.3 (26.1)</td>
</tr>
<tr>
<td>Complex</td>
<td>29.0 (35.8)</td>
<td>50.0 (39.7)</td>
<td>56.0 (36.5)</td>
<td>76.0 (31.5)</td>
<td>52.8 (39.1)</td>
</tr>
<tr>
<td>Total posttest</td>
<td>7.4 (3.3)</td>
<td>7.7 (3.4)</td>
<td>7.8 (3.9)</td>
<td>10.0 (2.8)</td>
<td>8.2 (3.5)</td>
</tr>
</tbody>
</table>

Note.  a Group 1 is the control group (no feedback; explain own solution only); b Group 2 received ambiguous feedback and explained both their own and an alternative solution; c Group 3 received feedback and explained the correct solution only; d Group 4 received feedback and explained both correct and incorrect solutions. Values enclosed in parentheses represent standard deviations. Scores by problem type are percent correct (0-100%). Total scores are number correct (0-14). N for each score by group is 20.

Various univariate, multivariate, repeated measure, and mixed model analyses if variances (ANOVAs) and covariances (ANCOVAs) were run, alpha levels were set at .05 and Pillai’s Trace F statistics and p-values are
reported for multivariate analysis results, unless otherwise noted.

**Covariates**

It was intended that several analyses of covariance (ANCOVA) with SAT scores as the covariates, would be used to analyze the experimental effects on total posttest scores, in order to control for existing aptitude differences. Correlations between SAT scores and posttest scores by problem type were expected to be unreliable, due to the nature of these measures. Total scores represent a broader range of knowledge, but scores by problem type represent more specific knowledge. Since SATs are used as indexes of general academic aptitude, it is reasonable to assume that total scores are more likely to correlate with them than are individual scores by problem type. As a result, SAT scores were not considered to be used as covariates in analyses involving scores by problem type.

Bivariate correlations between total pretest and posttest scores, and math and verbal SAT scores were calculated to verify the appropriateness of this method. These correlations between posttest and SAT scores were all found to be significant at the $p < .001$ level. Additional intended analyses included ANCOVAs with pre-test scores (both total and by problem type) as covariates. Bivariate
correlations between respective pre-test and post-test scores were also all found to be significant at the \( p < .001 \) level. These results confirm the appropriateness of including SAT scores and pre-test scores as covariates in analyses of posttest mean differences.

**Overall Practice Effects**

Overall practice session effects were tested by comparing total pre- and posttest scores across all groups and problem types, and by comparing pre- and posttest scores by problem type, across all groups. First, to determine whether the practice session had any general effect on posttest performance, a repeated measure (time) ANOVA was performed. Practice (time) was found to be a significant factor, \( F (1, 79) = 20.77, p < .001 \). As hypothesized, posttest scores were significantly higher than pretest scores. Mean scores for pre- and posttests were \( M = 6.71 \) (SD = 3.20) and \( M = 8.19 \) (SD = 3.48), respectively. (Maximum score = 14.)

Next, a repeated measures (time and Problem Type) MANOVA was used to examine problem type, and practice by problem type interaction effects. Again, practice was found to be significant, \( F (1, 79) = 17.69, p < .001 \), as was problem type, \( F (2, 78) = 93.32, p < .001 \). Not surprisingly, scores on simple-direct problems (\( M = 76.25 \),
$SD = 2.38$) were found to be significantly higher than scores on simple-indirect ($M = 46.13$, $SD = 3.00$) and complex ($M = 42.00$, $SD = 3.41$) problems. This analysis also revealed a practice by problem type interaction, $F(2, 78) = 10.72$, $p < .001$. Although there were no significant differences between pre- and posttest scores on simple-direct problems ($M_s = 76.56$ vs. $75.94$), there were slight improvements made (although not significant) on simple-indirect problems ($M_s = 42.00$ vs. $50.25$), and dramatic significant improvements made on complex test problems ($M_s = 31.25$ vs. $52.75$).

Practice session effects within each group condition were also analyzed by performing four individual (one for each group) repeated measure (time) ANOVAs, based on total test scores (maximum score=14). As expected, most students performed better on the posttest than on the pretest, however practice effects varied by group condition. Practice was found to be a significant factor for students in the Explain Correct and Explain Correct and Incorrect group conditions. The corresponding results were $F(1, 19) = 7.84$, $p < .05$ and $F(1, 19) = 14.63$, $p < .001$ for these groups, respectively. Table 1 illustrates means by group. The effect of practice approached significance for students
in the Ambiguous Feedback group, \( p = .09 \), but was not significant for students in the Control group.

Finally, practice effects were analyzed within each group and problem type, with the use of 12 individual repeated measure (time) ANOVAs. The results of these tests revealed significant practice effects for students in the Explain Correct, \( F (1, 19) = 17.79, \ p < .001 \), and Explain Correct and Incorrect, \( F (1, 19) = 14.14, \ p \leq .001 \), group conditions for complex problem type score differences. The means and standard deviations for pre- and posttest complex problem type scores for students in these groups are \( M = 20, \ SD = 19.47 \) and \( M = 56, \ SD = 36.48 \), and \( M = 44, \ SD = 41.85 \) and \( M = 76, \ SD = 31.52 \), respectively. Other combinations of group and problem type approached significance. Students in the Ambiguous Feedback group performed better on posttest than pretest scores for complex problem types, \( F (1, 19) = 3.62, \ p = .072 \), as did students in the Explain Correct group when solving simple-indirect problems, \( F (1, 19) = 3.35, \ p = .083 \), and students in the Explain Correct and Incorrect group when solving simple-direct problems, \( F (1, 19) = 3.35, \ p = .083 \). All mean scores for groups within problem type are included in Table 1.
**Group Effects**

Next, the various effect that each practice session, or group condition had on improved performance from pretest to posttest was examined. Group effects were analyzed based on total scores and then on scores by problem type.

Partial correlations between Group and posttest scores, controlling for respective pretest and SAT scores, were calculated. Two of the four correlations examined were found to be significant. The partial correlation between Group and total posttest scores was $r (72) = .35, p \leq .003$, and the correlation between Group and posttest scores for complex problem types was $r (72) = .44, p \leq .001$. These correlations are support our predictions that experimental conditions associated with group assignment would affect posttest performance.

A one-way (Group) ANCOVA with total pretest and SAT scores as covariates, revealed that group condition had a significant effect on total posttest performance, $F (3, 70) = 3.12, p < .05$. As expected, improvements for students in each of the experimental group conditions exceeded those for students in the control condition. Students in the Explain Correct and Incorrect group outperformed all others, and students in the control group had the smallest
increase in performance between pre- and posttests. Posttest mean scores adjusted for covariates, are illustrated in Figure 1.

![Figure 1](image_url)

Figure 1. Total posttest mean scores by experimental group evaluated at common covariate values (Pretest = 6.65; SAT Math score = 585; SAT Verbal score = 583.) Posttests were scored from 0 to 14. Groups 1 through 4 refer to Control (n = 20), Ambiguous Feedback (n = 19), Explain Correct (n = 20), and Explain Correct and Incorrect (n = 18) groups, respectively. Posttest mean score for the Explain Correct and Incorrect group was significantly greater than the mean score for the Control group. All other pairwise comparisons were non significant.

Interaction effects between group and problem type were also analyzed based on a 3 (Problem Type) by 4 (Group) mixed model MANCOVA with pretest scores by problem type as covariates. Consistent with prior analyses, this procedure
revealed significant group, $F (3, 73) = 2.71, p \leq .05$, and problem type, $F (2, 72) = 3.45, p < .05$, effects. Interestingly, a group by problem type interaction was also found, $F (6, 146) = 3.16, p < .01$. The varying effect that group condition had on performance within each problem type is illustrated in Figure 2. Clearly, group condition had greatest effect on posttest improvement within complex problem type, and little effect within simple-indirect problem type scores. Students in the Explain Correct and Incorrect group exhibited a clear advantage over other students within complex problem performance.

As a result of the observed group by problem type interaction effect, group condition effects were examined within each problem type. Three individual one-way (Group) ANCOVAs with respective pretest scores (by Problem Type) as covariates revealed a significant group effect on posttest performance of complex problem types, $F (3, 75) = 5.36, p < .005$. Within the complex problem type, all group condition pairwise comparisons were reviewed. Again, as predicted, improvement in test scores from pretest to posttest was greatest for students in the Explain Correct and Incorrect group. Respective means, adjusted for covariates for the four groups were 31.33 (Control), 48.49 (Ambiguous Feedback), 62.16 (Explain Correct), and 69.02 (Explain
Correct and Incorrect). These results indicated significant posttest performance differences between students in the Control and Explain Correct and Incorrect groups, \( p < .001 \), the Control and Explain Correct groups, \( p < .005 \), and the Ambiguous Feedback and Explain Correct and Incorrect groups, \( p < .05 \).

Figure 2. Posttest mean scores by Group and Problem Type evaluated at common covariate values (simple-direct pretest score = 76.56; simple indirect pretest = 42.00; complex pretest = 31.25.) Posttests by problem type were scored from 0 to 100. Groups 1 through 4 refer to Control (\( n = 20 \)), Ambiguous Feedback (\( n = 20 \)), Explain Correct (\( n = 20 \)), and Explain Correct and Incorrect (\( n = 20 \)) groups, respectively.
Feedback Effect

In order to understand the effect that feedback had on problem solving, posttest mean scores for students in the Control group were compared to scores for students in the Explain Correct group. A one-way (Feedback Group) ANCOVA with total pretest and SAT scores as the covariates revealed a significant group effect, $F(1, 35) = 4.07$, $p \leq .05$. Students who received feedback outperformed those who did not. Posttest means for the Explain Correct and Control groups, adjusted for covariates are $M = 8.45$, $SD = 3.94$ and $M = 6.70$, $SD = 3.33$, respectively.

A group by problem type interaction was found with a 3 (Problem Type) by 2 (Feedback Group) mixed model (posttest scores by problem type) MANCOVA with pretest scores as the covariates, $F(2, 34) = 15.49$, $p < .001$. Figure 3 suggests that feedback had a much greater positive effect on complex problem performance than on other problem type performance. With regard to complex problem posttest scores, students who received feedback displayed a tremendous gain (pretest $M = 44$; posttest $M = 76$), while students who did not receive feedback demonstrated no improvement from pretest to posttest (pretest $M = 27$; posttest $M = 29$). A post-hoc one-way (Feedback Group) ANCOVA with complex problem type
pretest score as covariate provided evidence that the
difference reported above was significant, $F(1, 37) = 10.56$, $p < .005$. Significant feedback effects were not
found for simple-direct or simple-indirect problem type
posttest performance.

![Figure 3](image)

Figure 3. Posttest mean scores by Feedback Group and
Problem Type evaluated at common covariate values
(simple-direct pretest score = 76.25; simple-
indirect pretest score = 43.00; complex pretest
score = 23.50.) Posttests by problem type were
scored from 0 to 100. Problem Types 1, 2, and 3 refer to Simple-direct, Simple-indirect, and
Complex, respectively. The Control group ($n = 20$)
represented the “No feedback” condition and the
Explain Correct group ($n = 20$) represented the “Feedback” condition.
Self-Explanation Effects

Posttest mean scores for students in the Explain Correct group were compared to scores for students in the Explain Correct and Incorrect group. A one-way (Self-Explanation Group) ANCOVA with total pretest and SAT scores as the covariates revealed a group effect that approached significance, $F(1, 35) = 3.49, p = .070$. Scores for students who self explained both correct and incorrect solutions did slightly better than students who only explained correct solutions. Total posttest means for the Explain Correct and Explain Correct and Incorrect groups, adjusted for covariates are $M = 9.33, SD = 2.81$ and $M = 8.40, SD = 3.94$, respectively.

There was no evidence of a group by problem type interaction based on a 3 (Problem Type) by 2 (Self-Explanation Group) mixed model (posttest scores by problem type) MANCOVA with pre-test scores as the covariates. However, a one-way (Self-Explanation Group) ANCOVA with simple-direct problem type pretest score as covariate revealed a significant group effect, $F(1, 37) = 6.59, p < .05$, as students who self-explained both correct and incorrect solutions ($M = 85.99, SD = 24.97$) outperformed students who explained only correct solutions ($M = 64.01$,
on simple-direct problem type posttests. No self-explanation group effect was found for simple-indirect or complex problem type posttest performance.

**Combined Effects of Feedback and Self-Explanation**

To examine the combined effects of experimental feedback and self-explanation conditions, the performance differences between students in the Control group and students in the Explain Correct and Incorrect group, were compared. First, a one-way (Group) ANCOVA with total pretest and SAT scores as covariates, was performed to determine combined effects on total posttest performance. As predicted, a significant group effect was present, $F(1, 33) = 11.73$, $p < .005$. Students in the Explain Correct and Incorrect group ($M = 10.00$, $SD = 2.81$) displayed much greater improvement in posttest scores than did students in the Control group ($M = 7.35$, $SD = 3.33$).

A 3 (Problem Type) by 2 (Group) mixed model MANCOVA with pretest scores by problem type as the covariate was used to compare combined experimental condition (Feedback and Self-Explanation) effects, and revealed a significant group by problem type interaction effect, $F(2, 34) = 5.29$, $p \leq .01$. As Figure 2 shows, combined experimental conditions (Group 4) had a much greater effect on complex problem type test performance than on other problem type performance. A
one-way (Group) ANCOVA with complex problem type pretest score as covariate revealed a significant combined condition effect, $F(1, 37) = 17.45$, $p < .001$, as students who were in the Explain Correct and Incorrect group ($M = 70.70$, $SD = 31.52$) strikingly outperformed students who were in the Control group ($M = 34.30$, $SD = 35.82$) on complex problem type posttests. No combined condition effect was found for simple-direct or simple-indirect problem type posttest performance.

Summary

Generally, results supported our predictions and hypotheses. The following summarizes these results by respective hypothesis.

H1: The directed practice session will lead to improved posttest performance (compared to pretest performance), in general.

Practice (or time) was found to be a significant factor. As expected, students on average performed better on the posttest than on the pretest. Practice was found to be a significant factor for students in the Explain Correct and Explain Correct and Incorrect group conditions, but not for students in the Ambiguous Feedback or Control groups. When practice effects within group and problem type were examined, significant practice effects were revealed for
students in the Explain Correct and Explain Correct and Incorrect group conditions for complex problem type score differences.

H2: Students in experimental groups will outperform students in the control group. Group was found to be a significant factor. Improvements for students in each of the experimental group conditions exceeded those for students in the control condition. Students in the Explain Correct and Incorrect group outperformed all others, and students in the Control group had the smallest increase in performance between pre- and posttests.

H3: Feedback will positively affect posttest performance.

Students who received feedback outperformed students who did not receive feedback, but only on complex problem type scores.

H4: The self-explanation of correct and incorrect solutions will positively affect posttest performance, in general.

No significant main effect of self-explanation was found, but one simple main effect was identified. Students who self explained both correct and incorrect solutions outperformed students who explained only correct solutions on simple-direct problem type posttests; however there were
no significant group differences on simple-indirect or complex problem types.

H5: Students who self explain both correct and incorrect solutions will demonstrate greater gains in performance from pretest to posttest on problems that require conceptual knowledge, than students who self explain correct (or “own”) solutions only. (i.e., The self-explanation experimental condition would have significant positive effects on simple-indirect problems.)

Results were not as anticipated. Experimental conditions positively affected performance on problems that could be solved with procedural knowledge (simple-direct and complex) however they did not seem to affect performance on problems that required more conceptual understanding (simple-indirect).

H6: The combined experimental condition would lead to significantly greater improvements in performance than the control conditions.

The combined effect of feedback and self-explanation condition was, of course, found to be significant. Students in the Explain Correct and Incorrect group displayed much greater improvement in posttest scores than did students in the Control group, especially with regard to complex problem types.
CHAPTER 4
DISCUSSION

Conclusions

The effects of feedback and self-explanation on learning were investigated, as measured by change in test performance from pretest to posttest. The effects of feedback and self-explanation have been examined separately under various conditions, and within various domains (Alibali, 1999; Chi et al., 1994; Ellis, 1997; Mwangi, & Sweller, 1998; Neuman, & Schwarz, 1998; Tudge et al., 1996). Since both have shown to have advantageous effects under many circumstances, they were used together in this study. Researchers have examined the effects that self-explanation has on students while learning information presented to them. To extend prior research, the self-explanation of correct as well as incorrect solutions was elicited (experimental condition) and compared to the condition in which only the correct answer was self-explained (control). It was expected that students who received feedback and who were asked to explain both correct and incorrect solutions would demonstrate the most improvement in test performance.
Various data analyses revealed many effects that were consistent with predictions and hypotheses. Significant main effects were found for practice. As expected, students performed better on the posttest than on the pretest, on average. More interestingly, group was also found to be a significant factor. Improvements for students in each of the experimental group conditions exceeded those for students in the control condition. Students in the Explain Correct and Incorrect group outperformed all others, and students in the control group had the smallest increase in performance between pre- and posttests. These results indicate that both feedback and self-explanation condition affected posttest performance.

As predicted, students who received feedback outperformed students who did not receive feedback. This significant posttest performance main effect (improvement) due to feedback is consistent with research by Tudge et al. (1996) and Ellis et al. (1995), and supports the assertion that feedback plays a significant role in cognitive development. The advantageous effects of feedback seem to be consistent across a wide range of ages. Tudge and others (1996) examined the impact of feedback on collaborative problem solving and found that the performance of 6- to 9-year-olds who received feedback during the problem solving
process improved significantly more than children who did not. In another study, Ellis et al. (1993) examined the effects of feedback and collaboration on the problem solving skills (decimal problems) of fifth graders, and found that while some of the children in the no feedback condition generated new incorrect strategies, only those in the feedback condition were able to generate new correct strategies. And, here we have found that feedback leads to greater improvements in problem solving performance of university students. Feedback, in terms of whether a response is correct or incorrect, seems to be a key factor in improving problem solving skills.

Chi and others have found that the spontaneous use of self-explanation (Chi et al., 1989; Ferguson-Hessler & de Jong, 1990) and also the elicitation of self-explanation (Chi et al., 1994) are associated with enhanced learning in the domains of physics and biology, respectively. However, Neuman and Schwarz (1998) concluded that only certain types of self-explanations, such as those that provide inferences or clarification, lead to improvements in problem solving. Clearly, additional research is necessary to determine exactly what it is about self-explanation that facilitates learning. Here, we attempted to extend existing knowledge by comparing students who were asked to explain correct
solutions to students who were asked to explain both correct and incorrect solutions.

Students who self explained both correct and incorrect solutions improved slightly more than those who explained correct only. This difference was not significant across all problem types, but it was found to be significant for posttest performance on simple-direct problem type scores. The nonsignificant main effect of the self-explanation condition studied could be the result of an insufficient sample size. This effect has not been studied in the past, therefore, it was difficult to approximate an effect size and ensure a sufficient sample size. Alternatively, perhaps the effect of self-explanation varies by problem type or task, which could explain the lack of significant effect found across all problem types. This claim would certainly be consistent with results from other studies (Mwangi & Sweller, 1998; Nathan et al., 1994).

An unexpected feedback condition by problem type interaction was also identified. Specifically, feedback was found to be a significant factor only on complex problem type scores. Students who self explained both correct and incorrect solutions outperformed students who explained only correct solutions on simple-direct problem type posttests, however there were no significant group
differences on simple-indirect or complex problem types. Neither group improved on simple-indirect problems, and both groups demonstrated significant improvement on complex problem type posttest scores. It seems reasonable to conclude that the information provided by feedback and self-explanations was insufficient for students to gain insight with regard to solving simple-indirect problems. On the other hand, the feedback offered to both self-explanation groups seemed to have provided enough information for students to realize significant gains on complex problems.

One explanation as to why the feedback provided might have been sufficient in the case of complex problems, but not in the case of simple-direct problems, has to do with the actual solutions provided by the experimenter during the directed practice session. That is, although the correct and incorrect solutions provided were the same for all students who originally furnished an incorrect or correct solution, respectively, the quality of those solutions may have been inconsistent across problem types. To explain, an example problem and the corresponding answers provided by the experimenter are shown for each problem type.
Simple-direct
Problem: Nancy is three-fourths Eileen’s height. Correct Solution: N = ¾ E. Incorrect Solution: E = ¾ N.

Simple-indirect
Problem: A large screen television costs 60% more than a medium screen television. Correct Solution: L = 1.6 M. Incorrect Solution: M = .6L

Complex
Problem: There are two boys for every three girls in Miss Anderson’s fourth-grade class. Correct Solution: 2/3 = b/g. Incorrect Solution: 2/3 = g/b.

The presentation of the simple-direct or simple-indirect solutions did not seem to afford any additional insight to students. However, the complex solutions provided did. For example, students often offered the equation “2b = 3g” for the problem above. When they were given the correct solution in the form “2/3 = b/g” many students demonstrated some level of recognition. There were many comments such as “Oh, yeah, that makes sense. It’s a ratio problem.” They often still did not exactly understand why their answer was incorrect, but they were able to recognize that the answer provided was better than their own. It is likely that if the correct solution had been provided in the form “2g = 3b”, this additional insight would not have been realized. It would be difficult to present a correct solution for simple-direct problems that would produce a similar effect, but it might be possible to
offer more informative solutions for simple-indirect problems. Consider the above simple-indirect problem. It is likely that a correct solution in the form “L = M + .6M” would have been more helpful than the solution actually provided. This type of manipulation is worth future consideration.

One explanation as to why equations presented in some forms might facilitate learning, while others may not is based on the concept of mental schemas. Sweller (1989) posits that schema acquisition is instrumental in learning, and that instructional techniques that facilitate schema acquisition will be more effective than techniques that do not. It seems plausible that students were more easily able to acquire correct mental schemas for the complex problems, given the form of the solutions presented, than for the other problem types. This idea could also help to explain why the instructional techniques used to teach mathematics in Asian classrooms is so successful. Educators in Japan will often present a new problem to students and allow them to develop strategies while working in small groups (Hatano & Inagaki, 1991, 1998; Stigler & Fernandez, 1995). Then, each strategy is examined and as a class students discuss the merits of each. Perhaps it is precisely the different information provided by various solution forms that
contributes to the knowledge that Asian students gain through classroom instruction. If we consider the fact that students come to a classroom with varying levels of knowledge and abilities, it is likely that equations presented in some forms may be more or less helpful in the acquisition of appropriate mental schemas. It follows then, that students might benefit from explaining several solutions, correct and incorrect.

Another unanticipated result was the dissimilar effect that the experimental conditions seemed to have on conceptual versus procedural problem solving. Experimental conditions positively affected performance on problems with solutions that were more dependent on procedural knowledge (simple-direct and complex); however, they did not seem to affect performance on problems that required a more conceptual understanding of algebra (simple-indirect). This is not consistent with conclusions offered by Nathan, Mertz, and Ryan (1994) who found that elicited self-explanation resulted in greater test improvement for algebraic tasks in which conceptual reasoning was necessary, but not for problems requiring only procedural equation manipulation. Although it is possible that these findings are contradictory, it is also possible that the disparity reflects a difference in problem classification.
(procedural versus conceptual). In their study, the tasks classified as procedural manipulation, were those in which the participant was instructed to directly solve for an unknown variable. The tasks intended to measure conceptual processing were those in which participants were instructed to express algebraically a provided story problem. In our study, all problem types required the interpretation of word problems or “stories” which, according to Nathan et al., does require some level of conceptual understanding. However, when each problem type is considered along with the respective solutions that were provided by the experimenter, it is reasonable to suggest that the simple-indirect problems require the highest level of conceptual understanding. Perhaps the simple-indirect problems were just too difficult.

Another possibility for this unexpected discrepancy can also be found in Nathan and others’ (1994) conclusions. In the same study discussed above, self-explanation had significant positive effects when students were in a low cognitive load condition, but not when students were in a high cognitive load condition. Nathan and colleagues suggest that the additional cognitive requirements of self-explanation might actually overtax students when the presented problems demand a high level of cognitive
resources. In our case, the simple-indirect problems require an additional algebraic manipulation which could easily translate to additional cognitive load.

This study provided valuable information with regard to the effects of feedback and self-explanation on students’ problem-solving performance, however, its design could be improved in several ways. First, in order to more effectively separate the effects of feedback and self-explanation, a group of students who explain both correct and incorrect solutions, but who receive no feedback is necessary. This would help us to determine whether the inclusion of feedback served to mitigate any self-explanation condition effects. Next, as mentioned earlier, the presentation of equations offered to students by the experimenter should be considered such that there are no differences in form by problem type, or such that the forms of equations provided are systematically manipulated. Another modification that would improve this study involves the assessment of conceptual versus procedural understanding. It is possible that all problem types require some level of conceptual understanding and that a better assessment would come from examining students’ verbalizations during the directed practice session. Finally, the varying results by problem type that were
observed may have been caused by differences in the level of cognitive resources demanded by the various problem types. This factor (cognitive load) should be consistent throughout, carefully measured and accounted for, or systematically manipulated.

These limitations notwithstanding, the results of this study extend our knowledge of the instruction strategies that could be used to facilitate the processes by which students acquire knowledge. Significant effects were observed for both self-explanation and feedback condition, however these effects varied across problem types. The evidence reveals that explicit feedback can be instrumental in helping students learn, especially with tasks such as the complex problem types presented here. Moreover, the self-explanations of correct and incorrect solutions led to significant improvements in performance when solving simple-direct problems. These findings offer insights that could prove valuable to educators, especially within the domain of algebra word problems.

**Suggestions for Future Research**

Considerable data from this study still remain to be analyzed. Findings related to the effects of feedback and self-explanation on strategy generation, modification, and selection during directed practice could offer important
insights. A main effect of feedback but not self-explanation on strategy generation may be interpreted as support for the associative models of these mechanisms, whereas a main effect of self-explanation may be interpreted as support for the combined model proposed by Crowley et al. (1997).

Perhaps we could use verbalized strategies to measure conceptual understanding, rather than rely on problem type. Conceptual understand as demonstrated through strategy verbalizations should also be compared to posttest performance. The relationship between procedural and conceptual knowledge has yet to be fully understood. A main effect of self-explanation on strategy generation and selection but not on posttest performance would lend support to Nathan and colleague’s (1994) suggestion that self-explanation affects the acquisition of conceptual but not procedural knowledge (i.e., assuming that pre- and posttests are better measures of procedural knowledge).

It would also be interesting to analyze students’ certainty ratings to determine whether there are any correlations between uncertainty and strategy generation or adjustment. Increased uncertainty concurrent with strategy development as revealed in verbalized explanations would lend support for both Piaget’s concept of equilibration and
Crowley and colleagues' (1997) competitive negotiation model. It is expected that feedback will have an effect on uncertainty. Although it is probably less robust, it is reasonable to expect that self-explanation condition could also affect uncertainty levels.

Although most of the research reviewed that addressed effects of self-explanation on learning offers evidence that self-explanation is beneficial to students, Mwangi and Sweller (1998) and found that the elicitation of self-explanation had no significant effect on learning to solve "compare word" problems. The major difference between their study and those of others is that Mwangi and Sweller's participants were third graders, while other researchers have focused on older children (e.g., eighth graders) and young adults. Similar studies involving participants of varying ages will help to determine the effects that developmental differences and abilities have on the usefulness of educational tools such as external feedback and self-explanation. Cognitive load issues such as the ones suggested by Nathan and colleagues' (1994) work, should also be part of future studies across various age ranges, as it is likely that self-explanation itself requires more cognitive resources for younger children than for older students.
If we consider the Vygotskian (Vygotsky, 1978) model in which successful scaffolding occurs when an adult maintains interactions with a child within the child’s zone of proximal development, we can identify some factors that might influence the effects of self-explanation. While age is clearly a possible significant factor, others include individual aptitude differences, problem difficulty, and cognitive load. More can be learned about possible effective uses of self-explanation by systematically manipulating and/or controlling these variables in future studies.

The goals of this study were to extend our knowledge of the mechanisms by which students acquire knowledge and the instructional strategies that could be used to facilitate these processes. Further research should be conducted in this area in order to determine which techniques are most effective under which circumstances. Findings offer insights that could prove valuable to educators in selecting age and task appropriate instructional strategies.
LIST OF REFERENCES


BIOGRAPHICAL SKETCH

Laura A. Curry was born on March 28, 1964, in West Islip, New York. She went to Sachem High School on Long Island and went on to receive a Bachelor of Arts in Economics from Bucknell University in 1986. After college, Laura embarked upon a career in actuarial science. She attained designations of Enrolled Actuary and Fellow of the Society of Pension Actuaries, and established herself as a pension consultant. Although this career was suitably challenging and afforded Laura the opportunity to capitalize on her mathematical strengths, she desired a career that was more personally fulfilling. In 1998, she entered the graduate program in developmental psychology at the University of Florida. Here, she is able to use her interests in statistical analysis towards the study of cognitive development. Laura’s research interests include the development and employment of problem-solving and decision-making skills during adolescence, and the evaluation of statistical methods used in social science research.