UNIVERSITY OF ALBERTA

ANATOMY OF A SHORTCUT BASED ON THE PRINCIPLE OF INVERSION

BY ©

KATHERINE MACLEOD ROBINSON

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE.

DEPARTMENT OF PSYCHOLOGY

Edmonton, Alberta

FALL, 1993

National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-315-88426-6

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Katherine M. Robinson

TITLE OF THESIS: The Anatomy of a Shortcut Based on the Principle of Inversion

DEGREE: Master of Science

YEAR THIS DEGREE GRANTED: Fall, 1993

10726-67 Avenue
Edmonton, Alberta
Canada, T6H 1Z9

July 15, 1993

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled ANATOMY OF A SHORTCUT BASED ON THE PRINCIPLE OF INVERSION submitted by KATHERINE M. ROBINSON in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE.

Jeffrey Bisanz

C. Donald Heth

Thomas Kieren

Date: ____________

Abstract

The purpose of the present experiment was to examine (a) age differences in students' spontaneous use of a shortcut based on the principle of inversion, (b) precursors of discovery and generalization of the shortcut, (c) stability of shortcut use over time, and (d) veridicality of self-report data. First, 40 students in Grade 1 and 40 students in Grade 5 were given a problem-solving task comprised of three-term problems. Although no age differences were found in spontaneous use of the inversion-based shortcut, different reasons for shortcut use were hypothesized to exist for students in each grade. A subset of students who did not spontaneously use the shortcut participated in a microgenetic study. Students were involved in seven problem-solving sessions and the precursors of discovery and generalization of the shortcut were investigated. Precursors were found to be different for students in each grade. Exposure to a task in which students were asked to recognize and justify appropriate shortcut use was sufficient to promote discovery and generalization for Grade 5 students. Direct instruction of the shortcut was necessary for discovery and generalization for Grade 1 students who had not yet discovered the shortcut. Finally, students who spontaneously used the shortcut frequently or

infrequently were assessed two months later to examine whether shortcut use remained stable over time. The use of the shortcut increased slightly between beginning and end of the experiment, due to an increase of shortcut use by the infrequent users. Although no age differences were found in spontaneous use of a shortcut based on the principle of inversion, age differences were found in precursors of discovery and generalization, ability to recognize and justify appropriate shortcut use, and stability in shortcut use over time. Self-report data appeared to be veridical based on all of the results in this experiment.

Acknowledgements

Thank you to Jeff Bisanz for all the time, effort, and patience that he has invested in this thesis as well as for showing me the importance of being a thorough and dedicated researcher; to Drs. C. Don Heth and Thomas Kieren, members of my supervisory committee, for their insightful comments and questions; to Dale Stout who first interested me in children and arithmetic; to Don Mabbott (a.k.a. Dan Mobbatt) for breaking the trail for me on this thesis journey; to my parents, Ruby and Stuart, who have always believed in me and in getting a good education - it is primarily due to their support and encouragement that I have gotten this far; and finally, to my fiancé Christian Riegel, whose loving care and support were crucial in the completion of this thesis and because he is the best friend a person could ever have.

TABLE OF CONTENTS

LIST OF TABLES

Anatomy of a Shortcut Based on the Principle of Inversion

The cognitive development of arithmetic skills has been the focus of intense study in the last decade. Several researchers have attempted to model how children select and learn arithmetic solution procedures (Ashcraft, 1987; Siegler & Shrager, 1984; Siegler & Shipley, 1992). Siegler, in particular, has emphasized the number and variety of solution procedures that children are capable of using to solve addition and subtraction problems. Siegler has shown that children are flexible in the use of solution procedures and that they know when to use them appropriately. In one study, Siegler (1987) found that 99% of all children used at least two solution procedures to solve the same types of problems and that 62% of all the children used at least three different solution procedures. These findings led Siegler to model the selection of solution procedures and to examine how new solution procedures fit into the model (Siegler & Shrager, 1984; Siegler & Shipley, 1992). What is unclear, however, is how children discover these new solution procedures. As Siegler noted, "new strategies do not emerge in a vacuum" (Siegler & Jenkins, 1989, p.7).

Children often learn new procedures to solve arithmetic problems on their own. However, studying how children discover new solution procedures to solve arithmetic problems is a difficult task. In most research the tendency has been to identify which solution procedures are used at different ages but to ignore how children learn or discover new solution procedures. This omission may be due to the lack of methodological approaches suitable for identifying transitional states in children's cognitive development. Despite this problem, the importance of examining the discovery of solution procedures should not be ignored. The study of development is, after all, the study of change.

Developmental psychology has often ignored its own mandate of studying change (Siegler & Crowley, 1991). Determining how change occurs has been a very arduous task, mainly because of difficulties in finding an appropriate methodology. Traditionally, cross-sectional or longitudinal studies have been used to assess changes. Although cross-sectional studies are adequate for determining what children do at different ages, they yield inadequate information about how changes occur because different children are assessed at different ages. Longitudinal studies typically involve such long intervals between testing sessions that the causes of change are unidentifiable. A third method, the microgenetic method, which is a type of longitudinal method, has been proposed as optimal for studying changes in cognitive development (Siegler & Crowley, 1991). The method is characterized by the intensive study of children throughout a period of change.

Siegler and Crowley (1991) described the microgenetic method as having three key properties. First, observations are taken from the beginning of the change being studied until the time a fairly stable state is achieved. Second, there is a high density of observation periods. Third, obtained data undergo detailed analyses in order to infer the qualitative and quantitative reasons for, and aspects of, the change. The method seems ideal for the study of development but is rarely used because the costs in terms of time and effort are prohibitive. As well, it is not always certain that the period of time in which the study is done will actually be the period of change for the subjects.

The microgenetic method has a long but rather obscure history (Catan, 1986). Werner was the first to use the term "microgenetic", although it seems that the method was previously used by another German researcher, Sander, at the beginning of the twentieth century. Sander conceived of Aktualgenese, as he called it, as a method for realizing developmental processes in the laboratory. Werner extended the use of Sander's methodology in an attempt to formulate general developmental laws that would apply to phenomena on all developmental levels. Vygotsky (1978) also advocated the use of the microgenetic method, as did Luria (1928). Despite the use of the microgenetic method by these prominent developmentalists, it still did not become widely used, probably because of its inherent difficulties.

Siegler and Crowley (1991), in a recent study, successfully examined the discovery of a new solution procedure in arithmetic. They used the microgenetic method to examine how children discovered the min solution procedure on addition problems. For a problem such as 2 + 9, instead of starting with the number 2 and adding 9 more, the min solution procedure is to start with the largest number, 9, and add the smaller number, 2, on to it. Contrary to the common belief that the discovery of a new solution procedure typically follows failures (Siegler, 1976), Siegler and Crowley found that successes as well as failures were likely to precede discovery of new solution procedures. Moreover, although learning models often include the assumption that impasses (difficult problems) precede new discoveries (e.g., VanLehn, 1988), Siegler and Crowley found that discovery of the min solution procedure was independent of the ease or difficulty of a problem. In this study, precursors of discovery were extremely long latencies and unusual verbal protocols prior to the use of the new solution procedure (Siegler & Crowley, 1991). Difficult problems (impasse problems) seemed to promote the generalization of a new solution procedure rather than its discovery (Siegler & Jenkins, 1989). The microgenetic study yielded important insights about discovery of the min procedure that would not have been found without the microgenetic method. As Siegler and Crowley stated: "Only by densely sampling strategy use prior to and during construction of the new approach could we have learned anything about which components presented the final obstacles to the discovery, about improvements in existing strategies that may have made possible construction of an essential component within the new strategy, or about children never using illegitimate strategies" (p. 615).

Correct usage of a solution procedure does not necessarily imply, however, that a child has explicit knowledge about the underlying concepts involved (Bisanz & Lefevre, 1990). Thus, children may have procedural knowledge in the absence of knowledge about the underlying concepts of a solution procedure, i.e., conceptual knowledge. This distinction has not been made in Siegler's model (Siegler & Shrager, 1984; and Siegler & Shipley, 1992), although Siegler and Crowley (1991) found, in their microgenetic study, that discovery of a new solution procedure seemed to require both conceptual and procedural knowledge.

Procedural and conceptual knowledge are different forms of understanding. The developing relation between procedural and conceptual knowledge is important for understanding how a new solution procedure develops. Therefore, a task that could be used to assess children's conceptual knowledge about underlying mathematical principles is necessary to evaluate children's understanding of mathematical principles. The microgenetic method could be a useful tool for identifying the mechanisms of change in the relation between children's understanding of procedural and conceptual knowledge. Assessing conceptual knowledge at various points in the study would yield information about the relation between procedural and conceptual knowledge. Of particular interest is whether conceptual knowledge follows, precedes, or is acquired at the same time as procedural knowledge.

Three-term arithmetic problems of the form $a + b - b$ or $b + a - b$ have some potential for assessing the relation between children's understanding of procedural and conceptual knowledge (Bisanz, Lefevre, & Gilliland, 1992; Dhaliwal, 1989; Starkey & Gelman, 1982). There are several ways to solve these problems successfully, but the most conceptually based (as well as the simplest) method is to use an inversion-based shortcut. This shortcut involves simply stating the answer a without performing any calculations. Because b is both added and subtracted, knowing that the answer is a can be deduced by simply looking at the problem carefully and understanding, in some sense, the arithmetic principle of inversion (i.e., that $a + b - b$ or $b + a - b$ must be equal to a).

Starkey and Gelman (1982) used three-term inversion problems of the form $a + b - b$ to assess 3- to 5-year-old children's understanding of the principle of inversion. They concluded that the children were capable of using the inversion-based shortcut to solve the problems. However, only accuracy measures were used and the children were only presented with one type of problem, which makes any conclusions about children's understanding of the principle questionable. Accuracy may have been just as high on three-term problems of the form a + b - c, or children simply may have stated the first term, a, as the answer without actually using a shortcut procedure. Bisanz, Lefevre, and Gilliland (1992) used the same three-term problems as Starkey and Gelman (1982) but added standard problems of the form a + b - c. If children used the principle of inversion to solve the inversion problems, the corresponding latencies would be significantly faster than latencies on standard problems. As well, two sizes of problems were used, small and large. If the inversion-based shortcut was being used, no differences in latencies should be found between small and large problems because the use of the shortcut is independent of the size of the numbers in the problem. Conversely, for small and large standard problems, latencies should increase as problem size increases because the larger the numbers in the problem, the longer it would take to calculate the answer. Bisanz et al. (1992) found that 6-year-olds and adults used the inversion shortcut but that 9-year-olds tended not to use the inversion shortcut. These conclusions were based on accuracy and latency data, which are more suitable for assessing procedural rather than conceptual knowledge. Accuracy and latency data revealed that some children probably used the inversion principle to help them solve the problems, but these data did not give any information about whether children knew why the inversion-based shortcut worked.

In a different study, Bisanz, Lefevre, and Gilliland (1989) found that 25 percent of 6-year-olds used inversion but that another 43 percent used negation (an intermediate strategy between successive addition and subtraction and inversion), 29 percent of both 7-year-olds and 9-year-olds used inversion, while 58 percent of 11-year-olds and 94 percent of adults used inversion. In a final study, Dhaliwal (1989) found that for Grade 2 (8-year-olds), 4 (10-year-olds), and 6 (12-year-olds) students, 36, 39, and 45 percent, respectively, of the students used either negation or inversion strategies, indicating that there was no significant difference in the use of inversion or negation strategies across grade. That at least some children of all ages are capable of using the principle of inversion to solve problems of the forms $a + b - b$ and $b + a - b$ seems clear. However, the age at which the inversion-based shortcut becomes the predominant solution procedure for solving inversion problems is unclear.

Rationale

The present experiment was designed (a) to examine the discovery of a shortcut based on the inversion principle and the precursors of that discovery, (b) to determine whether the shortcut generalizes to different problem type and size combinations and what the precursors of generalization are, and (c) to identify age-related differences in the use of the shortcut, in precursors of the shortcut discovery, and in the generalization of the use of shortcuts. To address these questions, the microgenetic method was used with stimuli consisting of inversion problems of the form $\underline{a} + \underline{b} - \underline{b}$ (simple) and $\underline{b} + \underline{a} - \underline{b}$ (reversed), and standard problems of the form $\underline{a} + \underline{b} - \underline{c}$. Because the inversion principle can be used to assess conceptual knowledge, knowing the mechanisms of change involved in the discovery of the inversion principle can yield important information about how conceptual knowledge is formed in arithmetic. This investigation is divided into three studies.

Study 1 was a pretest that preceded a microgenetic study. The pretest was used to identify solution procedures used by children in Grades 1 and 5. Grade 1 students were selected because it was hypothesized that the students had not had enough formal schooling to become rigid in their use of solution procedures. That is, they probably did not have enough practice with arithmetic problems such that they would always use the same rote procedures to solve them. Grade 5 students were selected because of previous findings that, by Grade 6, students solved inversion problems similarly to the way adults solved inversion problems. Thus, Grade 5 students would be almost ready to become "adult-like" in their solution procedures and how they changed was of interest.

This pretest was also done to identify students in these two grades who had not yet discovered the inversion-based shortcut. Most of these students were included in Study 2, the microgenetic study where they participated in a series of sessions and extra related tasks to identify the precursors of discovery as well as its generalization. Comparisons were made between the pretest and the posttest to assess changes in shortcut use. Study 3, the posttest, was done in order to evaluate the stability of the use of shortcuts over time for those students who had initially used the inversion-based shortcut either frequently or infrequently in Study 1.

Study 1

The first study was done to determine age-related differences in shortcut use as well as to identify the students who did not use the inversion-based shortcut and who would then form the experimental and control groups of Study 2. Based on previous studies, the percentage of students who used the inversion shortcut was expected to range from approximately 25 to 36% for Grade 1 students and from 45 to 58% for Grade 5 students (Bisanz et al., 1989; Dhaliwal, 1989). An age-related difference in shortcut use was expected (Bisanz et al., 1992). As well, the pretest was used to identify differences, if any, between inversion-based shortcut use on simple and reversed forms of inversion. Dhaliwal (1989) found that the shortcut was used more frequently on inversion problems of the form $a + b - b$ than on problems of the form $b + a - b$. This finding is not surprising because in simple inversion problems the *b*s are adjacent, whereas in the reversed form of inversion problems the *b*s are separated. In contrast to previous studies, three types of data were collected in this study: accuracy, latency, and self-report data. The veridicality of self-reports could therefore be assessed using the accuracy and latency data.

The importance of using self-report data in conjunction with accuracy and latency measures should be emphasized (see Siegler, 1987, 1989), especially in tasks that are used to examine the relation between conceptual and procedural knowledge. Siegler (1987, 1989) found that using self-report data can reveal information about how children solve arithmetic problems that would never have been found by only using accuracy and latency data.

Results of previous studies indicate that at least some young children use an inversion-based shortcut to solve inversion problems (Bisanz et al., 1989; 1992; Dhaliwal, 1989; Starkey & Gelman, 1982). Young children may be using a shortcut because they are not competent at retrieving or counting large numbers and are therefore forced to look for more innovative solution strategies.

In the present study, 40 students in both Grades 1 and 5 were given a set of three-term problems to solve. The problems were of three types and two sizes. Two types of inversion problems were used: simple inversion problems of the form $a + b - b$; and reversed inversion problems of the form $b + a - b$. A third type consisted of standard problems of the form $a + b - c$, which were included to assess whether students were using an inversion-based shortcut on inversion problems. If a left-to-right calculational strategy was being used on all three types of problems, then there would be no significant differences in solution latencies. If, however, a shortcut was being used on the inversion problems, then latencies would be faster than on standard problems where no shortcut was being used. The two sizes of problems were small (e.g., 2 + 7 - 7) and large (e.g., 3 + 24 - 24). As problem size increases, accuracy will tend to decrease and latency will tend to increase because of the greater calculational difficulties with large problems as compared to small problems. This change in accuracy and latency depending on the size of the problem is known as the problem-size effect (Koshmider & Ashcraft, 1991). However, if students use a shortcut, problem size should be irrelevant because no computation is being done. As in previous studies, accuracy and latency measures were recorded but a larger emphasis was placed on self-reports and observations of strategy use. This emphasis is based on the finding that chronometric data can obscure some important findings that are revealed with the use of self-reports (Siegler, 1987, 1989).

Several questions were investigated in this study.

Because the inversion-based shortcut is a good measure of children's underlying knowledge about the principles of arithmetic, we were interested in age differences in shortcut use. We wanted to know how many students used the inversion shortcut, how frequently they used it, and on what type and size of inversion problems students were most likely to use this shortcut. We looked at whether there was a significant difference between simple and reversed inversion problems to see whether the order of the three terms (ie., $\underline{a} + \underline{b} - \underline{b}$ vs. $\underline{b} + \underline{a} - \underline{b}$) in the problem affected shortcut use. Finally, we looked at how well accuracy and latency measures predicted shortcut use compared to self-report data in order to assess whether the findings of previous studies were supported.

As well as examining the use of inversion-based shortcuts, we were interested in whether the self-report data would reveal other, unexpected, solution procedures for solving three-term problems. Self-report has not been used extensively in the investigation of the inversion-based shortcut use, so other shortcuts may be used by the students that have not yet been reported. We examined which of the solution procedures tended to be most accurate and least accurate, as well as which of the procedures were the most rapid or slow.

Method

Subjects

Forty students in each of two grade levels, 1 and 5, participated in the study. Within each grade, half of the students were male and half were female. Median ages for the two grades were (in years:months) 6:10 (range 6:2 to 7:8) and 10:11 (10:2 to 11:10), respectively.

Materials

The stimulus list was comprised of six simple inversion, six reversed inversion, and twelve standard problems. Simple inversion problems were of the form $a + b - b$ (where $a <> b$), and included three small problems ($1 < a,b < 10$) and three large problems ($1 < a < 10$, and $20 < b < 30$). Reversed inversion problems were of the form $b + a - b$ (where $a <> b$), and also included three small and three large problems. Standard problems were of the form $a + b - c$ (where $a <> b$ or c and $b <> c$), and included six small problems ($1 < a,b,c < 10$) and six large problems ($1 < a < 10$, and $20 < b,c < 30$). Standard problems were approximately matched to inversion problems in that the sum of $b + c$ in standard problems was approximately equal to $2b$ in the corresponding inversion problems.

If a counting-based procedure was used on both inversion and standard problems, the amount of counting would be roughly equal and therefore the solution latencies and accuracies would be roughly the same for all types of problems. The 24 problems were ordered unsystematically with the constraints that no two problems of the same type and size were presented consecutively, no more than two inversion problems of either type or two standard problems were presented consecutively, and a small inversion problem (of either type) was always presented first and last. Two list orders were used, one the reverse of the other (Appendix A). All problems were presented on 21 X 28 cm white paper with a transparent plastic cover, with a maximum of six problems per page.

Procedure

All subjects were tested individually in one session that lasted approximately 30 min for Grade 1 and approximately 20 min for Grade 5 students. Students were given a preliminary briefing (Appendix B) and practice problems (Appendix C) before starting the session.

Students were asked to solve a set of 24 problems. Grade, sex, and problem order were counterbalanced. Each page was covered by a dark piece of paper and problems were uncovered one at a time. As soon as a problem was uncovered and the student attended to it, a hand-held stopwatch was started. Students were asked to give the answer and then to explain how they solved the problem. Latency data for answers and explanations were recorded on data sheets. If students were unable to solve the problem or had difficulty with a problem, then a discontinuation procedure was implemented after 40 s (Appendix D). The discontinuation procedure involved asking how the student was trying to solve the problem or how the problem could be solved. After the Grade 1 students had finished solving the 24 problems, they were given two simple two-term problems so that they were sure to end the testing session on a positive note. Grade 5 students were expected to have little difficulty with the experimental problems and were therefore not given these two extra problems.

Results and Discussion

Several analyses of variance with repeated measures were performed on accuracy, latency, and self-report data. Accuracy and latency data were analyzed to determine whether these measures are consistent with the notion that students might have used a shortcut on inversion problems, as they reported.

<u>Accuracy</u>

Accuracy measures were based on the first response that students gave after they had solved a problem on the assumption that the first response accurately reflected the procedure students used initially to solve a given problem. In contrast, final responses different from the initial responses could reflect changes in solution procedures, possibly affected by experimenter prompts, and as such might not reflect the procedure that the students had originally used to solve the problem. In the analyses of accuracy, proportions of accuracy were calculated on correct responses before the discontinuation protocol. On problems where the discontinuation protocol was implemented, the solution was considered to be inaccurate. Percentages of discontinuation protocols that were used as a function of grade, problem type, and problem size are included in Table 1. Grade 5 students hardly ever had a discontinuation protocol implemented. However, Grade 1 students had many discontinuation protocols, implying that these problems were very difficult for the students to solve, especially when the problems were large. Percentages of changed answers are included in Table 2. Grade 1 and 5 students changed their answers infrequently.

Insert Table 1 about here

Insert Table 2 about here

Analyses of accuracy were done in order to assess whether older students were more accurate than younger students, whether small inversion problems were more accurate than large, and whether inversion problems were more accurate than standard problems. Finding that small and large inversion problems were answered with equal accuracy would be consistent with the notion that a shortcut is being used. If inversion problems were answered more accurately than standards, then this result also would be consistent with the use of the inversion-based shortcut.

Proportions of correct responses were subjected to a 2(Grade: 1 and 5) x 2(Sex: male and female) x 2(Size: small and large) x 3(Type: simple inversion, reversed inversion, and standard) analysis of variance with repeated measures on the last two variables. Grade 5 students were more accurate than Grade 1 students (.92 vs. .44), $F(1,76) = 114.62$, $p < .01$, and students were more accurate on small than large problems (.73 vs.

.44), $F(1,76) = 23.77$, $p <$ 01. Accuracy varied as a function of problem type, $F(2,152) = 47.46$, $p < .01$. Tests of simple effects revealed that students were more accurate on simple and reversed inversion problems (.76 and .75) than on standard problems (.54) ($p < .01$).

These main effects are qualified by three interactions. First, grade interacted with size, $F(1,76) = 9.42$, $p < .01$ (see Table 3). Tests of simple effects revealed that Grade 5 students were more accurate than Grade 1 students at both levels of size ($ps < .01$), and that large problems were more difficult than small problems for the younger students ($p < .01$) but not for older students. A second interaction was found between grade and type, $F(2,152) = 20.39$, $p < .01$ (see Table 3). Although students in both Grades 1 and 5 were more accurate on inversion problems than standard problems, the difference in accuracy between the inversion and standard problems is larger in Grade 1 than Grade 5. These findings must be interpreted with the consideration that there may have been ceiling effects: Older students were almost perfectly accurate, especially on inversion problems. The final interaction was between type and size, $F(2,152) = 7.43$, $p < .01$. Although there was no size effect for simple inversion (.77 and .74 for small and large, respectively), students solved small problems more accurately than large problems for both reversed inversion (.80 and .71) and standard problems (.62 and .45) (ps < .01).

Insert Table 3 about here

The main effects of grade and size are consistent with findings that older students are generally more accurate and small problems are easier to solve than large problems. That both types of inversion problems are more accurate than standard problems is consistent with the use of a shortcut on inversion problems. The lack of any significant difference between both types of inversion problems may indicate that an inversion-based shortcut was used equally on both types of problems.

The interaction of grade and size reflects the ease that Grade 5 students had with both the small and large problems and the greater difficulty that Grade 1 students had with small and, especially, large problems. For simple inversion there was no problem-size effect, indicating that the inversion-based shortcut may have been used, at the very least, on the large problems. That students were less accurate on large than small problems for reverse inversion and standard problems indicates that perhaps shortcuts were used less frequently on these two types of problems. The interaction between grade and type indicates that, although students were more accurate on inversion problems than standards at both grades, the problem type effect was larger in Grade 1 than in Grade 5. However, the interaction may be a by-product of ceiling-level performance in the older students.

<u>Latency</u>

Latency measures were calculated as the median response time for each combination of problem type and size. Only correct responses were included in the analyses. Because younger students had such a high frequency of discontinuation protocols (48.65% overall), especially on large standard problems, discontinuation protocols were included in the median calculations as 40 s on the premise that the younger students may have solved the problem correctly if given more time. In view of this, latency data from Grade 1 students are assumed to be underestimates of true values and should be interpreted cautiously.

The latency analyses were done to assess whether small inversion problems were solved faster than large problems and whether inversion problems were answered faster than standard problems. If responses to small and large inversion problems were equally fast, then these data would be consistent with the notion that a shortcut is being used: With the inversion-based shortcut no calculations are necessary, and thus latencies should be unaffected by problem size. Finding that inversion problems were solved faster than standards also would be consistent with use of the inversion-based shortcut.

Separate latency analyses were done for Grades 1 and 5. In an overall analysis the larger variability in Grade 1 latencies might obscure the smaller but still significant differences among means in Grade 5. Grade 1 students ranged from being very fast to having a maximum latency (40 s), whereas Grade 5 students had a much smaller range of latencies. Latencies for each problem type and size combination were based on the median of correct initial responses.

Medians of the correct response latencies for each grade were subjected to a 2(Sex: male and female) x 2(Size: small and large) x 3(Type: simple inversion, reversed inversion, and standard) analysis of variance with repeated measures on size and type. Analyses at both grades revealed the same main effects and interaction. Both Grade 1 and Grade 5 students were faster on small than large problems (25.2 s vs. 29.4 s, $F(1,36) = 13.86$, $p < .01$, and 4.6 s vs. 6.0 s, $F(1,38) = 8.45$, $p < .01$, respectively). Students in Grade 1 were faster on both types of inversion problems than on standard problems (24.2 s, 22.7 s, and 35.0 s), $F(2,72) = 30.51$, $p < .01$. The same result was true for Grade 5 students (4.3 s, 4.9 s, and 6.6 s), $F(2,76) = 18.45$. $p < .01$. There were no significant differences between the inversion problem types for either grade.

These main effects of type and size are qualified by an interaction of type and size for both grades, $F(2,72) = 4.38$, $p < .05$, and $F(2,76) = 8.95$, $p < .01$, respectively (see Table 4). For Grade 1, although there was no size effect for simple inversion, students had shorter latencies on small problems than large problems for both reversed inversion and standard problems ($ps < .01$). Further analysis revealed the increase in latency from small to large was significantly greater for standard problems than reversed inversion problems. For Grade 5, there was no size effect for either simple or reversed inversion problems, although latencies were shorter on small problems than large problems for the standards.

Insert Table 4 about here

The main effects of size are consistent with findings that small problems are easier to solve than large problems. That both types of inversion problems are faster than standard problems is consistent with the use of a shortcut solution procedure on inversion problems. The lack of any significant difference between both types of inversion problems may indicate that inversion is being used equally on both types of problems.

For Grade 1 students there was no problem-size effect for simple inversion problems, indicating that the inversion-based shortcut was being used at least on large problems. That the Grade 1 students were faster on small than large problems for reverse inversion and standard problems indicates that perhaps there was less frequent use of shortcuts on these two types of problems. However, there was a greater size effect for standard than reversed inversion problems, which could indicate that although there was less shortcut use on reversed than simple problems, there was still some shortcut use.

Grade 5 students were faster on both types of inversion problems than on standard problems when they were large, but all three problem types were equally rapid for small problems. This pattern of data is consistent with the use of a shortcut on large inversion problems, but the lack of a difference between small inversion and standard problems indicates that retrieval is just as fast as shortcut use when the problem size is small.

All of the latency results are consistent with the accuracy results and provide two sets of evidence for the use of the inversion-based shortcut by both Grade 1 and 5 students.

Self-Reports and Observations

Self-report data consisted of the responses students gave to the question: "How did you solve this problem?" Their responses were categorized and analyzed to identify the understanding of underlying arithmetic principles and the procedural knowledge that children had about the problems. When no contradictory behavioural data were observed (e.g., the student counted on his fingers to solve the problem but reported solving the problem in his head), the students' self-reports were used. When self-report and behavioural data were discrepant, the behavioural data were used to classify the solution procedure. Such inconsistencies were infrequent.

Self-report data were classified into six solution procedure categories. The <u>left-to-right solution procedure</u> involved addition of the first two numbers in the problem and then subtraction of the last number from the sum [($\underline{a}$ + $\underline{b}$) - $\underline{c}$]. For example, for the problem 4 + 3 - 5, a student might say, "I added the 4 and the 3 and got 7 and then I took away 5 and that left me with 2." The <u>subtraction-first shortcut</u>, in contrast, involved subtracting the third number from the second and then adding the first number [$\underline{a}$ + ($\underline{b}$ - $\underline{c}$)] for standard problems. For example, for the problem 8 + 23 - 21, a student might say, "I took 21 away from the 23 and that left 2 and 2 + 8 is 10 so the answer is 10." In <u>negation</u>, students added the first two numbers and then simply restated the first number as if the student had realized that the number to be subtracted had just been added. For example, for the problem 22 + 7 - 22, a student might say, "I added the 22 and the 7 together and got 29 but then 22 take away 22 is 0 so the answer is just 7." The <u>inversion-based shortcut</u> was used when the $\underline{a}$ was simply stated as the answer because the $\underline{b}$ terms were the same and cancelled each other out [$\underline{a}$ + ($\underline{b}$ - $\underline{b}$) = $\underline{a}$]. For example, for the problem 9 + 27 - 27, a student might say, "the 27 takes away the other 27 so only the 9 is left." _Other_ was used for solution procedures that were used very infrequently but the students reported a solution procedure that, although not necessarily giving the correct answer, was understandable. For example, for the problem 5 + 8 - 4, a student might say, "the answer is 5 because 5 is the first number." _Ambiguous_ was used for any self-report solution procedure that could not be identified clearly. For example, for the problem 4 + 8 - 6, a student might say, "you add 4 + 8 and take 3 from the 8 and put it on the 4 and that gives you 9."

Proportions of these categories are presented as a function of grade, problem type, and size for Grade 1 students in Table 5 and for Grade 5 students in Table 6. Reliabilities for each of these six solution procedure categories were calculated for the same set of problems administered at a later time on a subset of 12 subjects. These reliabilities are presented in Table 7 as a function of problem type and size. Reliabilities were calculated as a proportion of agreements between two raters out of the total number of agreements and disagreements. Reliabilities were high for frequently occurring shortcuts and procedures but much lower for infrequently occurring categories.

Use of two shortcuts was examined in detail: the inversion-based shortcut on inversion problems and the subtraction-first shortcut on standard problems. Examining the left-to-right strategy would have yielded little new information because the sum of the relative frequencies for the shortcuts and the left-to-right strategy nearly totalled 100%.

Insert Table 5 about here

Insert Table 6 about here

Insert Table 7 about here

Inversion-based shortcut. Proportions of self-reported shortcuts on inversion problems were subjected to a 2(Grade: 1 and 5) x 2(Sex: male and female) x 2(Size: small and large) x 2(Type: simple inversion and reversed inversion) analysis of variance with repeated measures on the last two variables. Standard problems were excluded because the inversion-based shortcut could not be used appropriately on these problems. However, although Grade 5 students never used the inversion-based shortcut on standard problems, one Grade 1 student used the shortcut on 8 out of 12 standard problems, indicating an overgeneralization of the shortcut.

The analysis of self-reports was done in order to see whether there were any age differences in students' self-reports of the inversion-based shortcut. It was expected that as age increases, so would shortcut use, based on the results of earlier studies. The analyses were also done to see whether the results of the accuracy and latency analyses were also paralleled by the self-report analyses.

Interestingly, there was no main effect of grade, indicating that students in both grades reported using the shortcut equally, although older students reported using the inversion-based shortcut slightly more often (.39 vs. .44). Students reported more shortcut use on large than small problems (.48 vs. .34), $F(1,76) = 45.41$, $p < .01$. Students also reported more shortcut use on simple than reversed inversion problems (.44 vs. .38), $F(1,76) = 4.54$, $p < .05$. These main effects are qualified by two interactions, the second of which is included in the first. The first was the four-way interaction between grade, sex, type, and size $F(1,76) = 7.21$, $p < .01$ and the second interaction was between grade, sex, and size $F(1,76) = 3.97$, $p < .05$ (see Table 8). Grade 1 males used the shortcut more frequently on large than small problems for the simple type, but there was no size effect on reversed inversion problems. There was no type effect for small problems but Grade 1 males used the shortcut more on simple than reversed for large problems. Grade 1 females and Grade 5 males had no size effect for simple inversion problems but used the shortcut more on large than small problems for reversed inversion problems. There was no type effect for either size for Grade 1 females and Grade 5 males. Grade 5 females used the shortcut more on large than small problems for both types of inversion problems. There was no type effect for small problems but Grade 5 females used the shortcut more on the simple than the reversed type for large problems.

Insert Table 8 about here

The finding that there was no main effect of grade is surprising, considering the results of previous studies. However, upon reflection, perhaps the results should not be so surprising. We know that at least some younger students are capable of using an inversion-based shortcut, and we also know that younger students have difficulties retrieving as well as counting large numbers and therefore must look for other ways to solve the problems, such as the inversion-based shortcut. Older students are capable of using retrieval, have more practice with large problems, and consequently may not be looking for a shortcut to solve the problems more easily or more quickly. Therefore, not all of the older students might notice that the shortcut can be used.

The main effect of size is consistent with the use of a shortcut on harder problems, i.e., problems that involve large numbers. The difference in shortcut use between the two types of inversion problems implies that adjacency of the b terms makes the possibility of using the shortcut somewhat more obvious than when the b terms are nonadjacent, although the difference is not large.

Subtraction-first shortcut. Just as the inversion-based shortcut is effective for solving inversion problems, the subtraction-first shortcut is an effective shortcut for solving standard problems. Subtracting the third number from the second ($b - c = d$) and then adding the difference to the first number ($d + a$) requires much less computation than a left-to-right strategy and would make solving large problems almost as easy as small problems. Although this shortcut was not nearly so prevalent as the inversion-based shortcut (.11 vs. .41), it was still used often enough to be of interest. When using the subtraction-first shortcut there can be some difficulties in that the second number is not always larger than the third, resulting in a negative difference. For Grade 1 students, when the difference was positive, accuracy was .64. When the difference was negative, accuracy dropped to .34. Grade 5 students had a much higher accuracy rate but still had a large discrepancy between problems with positive and negative differences, .99 vs. .46. The analysis of the subtraction-first shortcut was done in order to assess whether there were any age differences and whether the shortcut was used differentially on small and large problems.

Proportions of subtraction-first use on standard problems was subjected to a 2(Grade: 1 and 5) x 2(Sex: male and female) x 2(Size: small and large) analysis of variance with repeated measures on size. Grade 5 students reported using the subtraction-first shortcut more often than Grade 1 (.17 vs. .04), $\underline{F}(1,76) = 8.62$, $\underline{p} < .01$. Students also reported using the subtraction-first shortcut more on large than small problems (.13

vs. .09), $F(1,76) = 14.05$, $p < .01$. These main effects were qualified by a grade by size interaction, $F(1,76) = 6.25$, $p < .05$ (see Table 9). Grade 5 students reported using the subtraction-first shortcut more on large than small problems but there was no problem-size effect for Grade 1 students. These results need to be interpreted cautiously because of floor effects.

Insert Table 9 about here

The finding that Grade 5 students use the subtraction-first shortcut more than Grade 1 students may be the result of older student's greater understanding of the principle of associativity. The main effects of size are consistent with the use of a shortcut on harder problems, i.e., problems that involve large numbers.

The interaction between grade and size seems consistent with students using a shortcut more on harder problems. For Grade 5 students, harder problems were the large problems, whereas the Grade 1 students found both the small and large standard problems difficult.

Correlations were computed to see whether there was a relation between reported inversion shortcut use and subtraction-first shortcut use. The frequencies of these two shortcuts were highly correlated in Grade 1 (r = .39, p < .001) and in Grade 5 (r = .60, p < .001), and also when the data from both grades were combined (r = .48, p < .001). Thus subjects who use the inversion-based shortcut more frequently are also more likely to use the subtraction-first shortcut. This finding raises the issue of individual differences in shortcut use.

Individual Differences

The analyses on accuracy, latency, and self-report data provide evidence for selective use of shortcuts but they obscure individual differences. To see whether there were consistent differences between students, self-report data were re-examined. Based on this examination, students were divided into three groups depending on how often they reported using an inversion-based shortcut. If the students self-reported using the shortcut on 75% or more of the 12 inversion problems, they were classified as Users (13 in Grade 1 and 14 in Grade 5). If the students said they never used the shortcut on any of the inversion problems, they were termed Non-users (13 in both grades). Students who said they used the inversion shortcut at least once but on less than 75% of the problems were classified as Variable Users (14 in Grade 1 and 13 in Grade 5).

Analyses of variance were recomputed on the accuracy and latency data with group as a between-subjects variable. All other independent variables remained the same. Only main effects and interactions involving group are reported. All other main effects and significant interactions in this second set of analyses were the same as those found in the first analyses. These analyses were done primarily to determine whether the students in each group were doing what they reported having done, that is, whether their self-reports were veridical. Therefore, Users should have much faster latencies and be more accurate than the Variable Users and Non-users, and the Variable Users would be faster and more accurate than the Non-users.

Accuracy. The re-analysis of accuracy data should show that Users would be more accurate overall than either Variable Users and Non-users and that Variable Users should be more accurate than Non-users. As well, Users should be more accurate on inversion than standard problems, whereas Non-users should be equally accurate on all types of problems. Variable Users should fall in between. There should be no problem-size effect for Users because they employ a shortcut on large and small problems. A problem-size effect should exist for Non-users and an intermediate problem-size effect should be found for Variable Users. If these results are found, then the self-report data would appear to be veridical.

Proportions of correct responses were subjected to a 2(Grade: 1 and 5) x 2(Sex: male and female) x 3(Group: user, variable user, and non-user) x 2(Size: small and large) x 3(Type: simple inversion, reversed inversion, and standard) analysis of variance with repeated measures on the last two variables. Users were more accurate than either Variable Users or Non-users (.82 vs. .56, .53), $\underline{F}(2,68) = 23.27$, $\underline{p} < .01$, with no significant difference between Variable Users and Non-users. These results must be interpreted cautiously due to ceiling effects for all of the Grade 5 students. In addition, Grade 1 Users were almost perfectly accurate on both types of inversion problems.

The main effect of group was qualified by three interactions involving group. The first interaction was between grade and group $\underline{F}(2,68) = 11.09$, $\underline{p} < .01$, and the second was between group and type $\underline{F}(4,136) = 21.41$, $\underline{p} < .01$. Both of these interactions were qualified by a third interaction between grade, group, and type $F(4,136) = 14.08$, $p < .01$ (see Table 10).

Insert Table 10 about here

For the Users, there was no age difference in accuracy for inversion problems (both types) but Grade 5 students were more accurate than Grade 1 students on standard problems. Grade 1 Users would still have difficulty solving standard problems for which they could not use the inversion-based shortcut. For the Variable Users and Non-users, Grade 5 students were more accurate than Grade 1 students regardless of problem type. Grade 5 students were almost perfectly accurate regardless of user group.

Grade 1 Users were more accurate than Variable Users and Non-users on both inversion type problems, while Variable Users were more accurate than Non-users. As well, there were no group differences for standard problems. The more inversion shortcut use by the Grade 1 students, the higher the accuracy on inversion problems. Grade 5 students, on the other hand, were equally accurate, regardless of group or problem type.

Grade 1 Users and Variable Users as well as Grade 5 Users were more accurate on inversion than standard problems with no difference between inversion types.

Grade 1 students had considerable difficulty with large problems if they did not use the shortcut. Any use of the inversion-based shortcut would dramatically increase their accuracy. Although Grade 5 students were highly accurate, most errors occurred on standard problems, where calculation errors could occur, and errors were almost nil on inversion problems for Users. Grade 1 Non-users as well as Grade 5 Variable Users and Non-users were equally accurate on all three types of problems. If the inversion-based shortcut was not being used or was used rarely, the inversion problems were no easier than the standard problems.

The main effect of group is consistent with the notion that the Users used the inversion-based shortcut, just as they reported. As for the interactions, the accuracy of students in Grade 1 on inversion problems corresponded to the user groups. Users were more accurate than Variable Users and Non-users, and Variable Users were more accurate than Non-users. On standard problems, group was irrelevant. Therefore, use of a shortcut appears to be unrelated to computational fluency. The accuracy of all Grade 5 students was the same, regardless of problem type or group and probably reflects ceiling effects. Grade 1 Users and Variable Users as well as Grade 5 Users were more accurate on inversion than standard problems, again indicating shortcut use whereas Grade 1 Non-users and Grade 5 Variable Users and Non-users were equally accurate on all problem types. The analysis of accuracy using group seems to indicate that students' self-reports were veridical.

Latency. The re-analysis of latency data should show that Users would be faster overall than either Variable Users and Non-users and that Variable Users are faster than Non-users. As well, Users should be faster on inversion than standard problems, whereas Non-users should be equally fast on all types of problems. Variable Users should fall in between. There should be no problem-size effect for Users because Users use a shortcut on large and small problems. A problem-size effect should exist for Non-users and an intermediate problem-size effect should be found for Variable Users. If the above results are found, then the veridicality of self-report data will be supported.

The latency analyses for Grades 1 and 5 were both 2(Sex: male and female) x 3(Group: users, variable users, and non-users) x 2(Size: small and large) x 3(Type: simple inversion, reversed inversion, and standard) analyses of variance with repeated measures on size and type. There was a main effect of group for both grades. In Grade 1, Users were faster than either Variable Users or Non-users (17.5 vs. 25.6, 37.4 s), $F(2,32) = 25.88$, $p < .01$ and Variable Users were faster than Non-users. In Grade 5, although there were no significant differences between Users and Variable Users, they were both faster than Non-users (4.0, 4.8, vs. 7.1 s), $F(2,34) = 5.60$, $p < .01$.

For Grade 1, the main effect of group was qualified by an interaction between group and problem type $F(4,64) = 25.64$, $p < .01$ (see Table 11). For both types of inversion problems, Users were faster than Variable Users and Non-users and Variable Users were faster than Non-users. For standard problems, group had no significant effect. In addition, Users and Variable Users were faster on inversion problems (simple and reversed) than standard problems, with no difference between inversion types. Non-users, although no faster on inversion than standard problems, were more rapid on reversed inversion than simple inversion problems. There were no interactions with group for Grade 5.

Insert Table 11 about here

The main effect of group in Grade 1 is consistent with the notion that the Users employed a shortcut based on inversion, just as they reported. Users were faster than Variable Users who were in turn faster than Non-users. In Grade 5, although there were no significant differences between Users and Variable Users, both groups were still faster than the Non-users, again indicating shortcut use. Although not all the problem-size effects were found that would be expected if students reports were veridical, Grade 5 students solved all problems quickly, even when using successive addition and subtraction. As well, the older students used a shortcut on the standard problems that would result in a lack of problem-size effect for this problem type. The interaction between group and type in Grade 1 indicates that the main effect of group holds for inversion problems but that group is irrelevant for standard problems, which would be expected because the shortcut cannot be properly used on standard problems. This finding is also consistent with usage of a shortcut being unrelated to computational fluency. That Users and Variable Users were faster on inversion than standard problems indicates the use of a shortcut on inversion problems. The finding that Non-users were faster on reversed inversion than simple inversion is unexpected.

Conclusions

The results of this study indicate that students in both Grades 1 and 5 are equally capable of using a shortcut based on the principle of inversion for solving three-term problems. The lack of age difference in shortcut use has never been found before. However, it may be that the reasons for students' shortcut use in each grade are different. The difficulty that Grade 1 students had in solving these problems may have promoted the use of an inversion-based shortcut because the problems would have been unsolvable otherwise. The Grade 5 students, however, had little difficulty with any of the problems and therefore may not have been looking for a shortcut to solve the problems. As well, it may be that schooling has an effect on the older students such that they are so used to solving problems with a left-to-right algorithm that they do not notice a shortcut as easily as the Grade 1 children did.

There was less inversion-based shortcut use on reversed inversion than simple inversion problems. As well, unlike simple inversion problems, there was more shortcut use on large than small problems. The size difference in shortcut use may be due to the relative ease of solving the small problems with a left-to-right solution procedure compared to the large problems. The order of the terms affects the use of shortcuts, and therefore data should not be aggregated across types of problems.

From this study, it can be concluded that self-reports are a good indicator of how children solve three-term problems. The accuracy and latency data closely parallel the self-report data and imply that the self-reports are reasonably veridical. Although the evidence for veridicality in Grade 5 is less compelling, lack of problem-size effects may be due to short latencies on all problems and the use of the subtraction-first shortcut on standard problems. Therefore, self-reports are a useful tool in the assessment of children's solution procedures for three-term arithmetic problems. Not only did self-report data parallel the accuracy and latency data, these data also revealed another shortcut that has never before been found in children. Indeed, if self-reports had not been used, the subtraction-first shortcut would not have been discovered. This finding of a new shortcut emphasizes the importance of using self-reports and the risk of obscuring important findings if they are not used.

Although students in both Grades 1 and 5 were equally likely to spontaneously use the inversion-based shortcut, age differences may exist in the reasons for use of the shortcut. The younger students who spontaneously used the shortcut may have had sufficient knowledge about the underlying concepts of arithmetic such that when presented with extremely difficult problems, were able to implement this conceptual knowledge to solve the problems by using a shortcut based on the principle of inversion. Therefore, when necessary, these students could employ their conceptual knowledge to solve difficult problems when solution procedures that they had been taught, such as successive addition and subtraction, could not be used. Older students, however, did not need to use the shortcut to solve problems, although the shortcut is a faster and easier solution procedure, especially on large problems. Therefore, the reasons for using a shortcut may reveal important age-related differences. Consistent individual differences were found within each age group. Why some students and not others spontaneously use the shortcut is a question whose answer would have important implications for development and instruction. As well, although younger students can use a shortcut, little can be inferred about the extent to which they understand the inversion principle. That is, although these students may use the shortcut, they may not exhibit other types of understanding, such as recognizing and justifying appropriate use of the shortcut. Again, there may be age related differences in the degree of understanding of the inversion principle.

Study 2

The second study, the microgenetic study, was done to investigate students' discovery of how to use an arithmetic principle, the principle of inversion, in solving three-term arithmetic problems. Students who never reported using an inversion-based shortcut to solve the inversion problems in Study 1 were assigned to experimental and control groups. These Non-users were chosen because we were interested in how students discover a new shortcut. The precursors of discovery, generalization of the shortcut, and age differences were all of interest. Information about the mechanisms of change involved in discovery, about what is needed for generalization of a shortcut, and about whether different precursors of discovery and generalization are needed for students of different ages would have important implications for the field of cognitive development as well as instruction.

As in the pretest, accuracy, latency, and verbal protocols (self-reports) were analyzed. The use of verbal protocols is often an important aspect of the microgenetic method (Catan, 1986). Despite concerns about the veridicality of verbal self-reports (Ashcraft, 1990; Cooney & Ladd, 1992; Ericsson & Simon, 1980), researchers have found that self-reports yield important information that would have remained undiscovered if not for verbal protocols (e.g., Siegler, 1987). Results of Study 1 also indicate that self-reports are reasonably veridical because of the close parallels among the latency, accuracy, and self-report data.

Students who never reported using a shortcut on the pretest were placed in either an experimental or control group. The experimental group had five weekly sessions in which they solved problems similar to those given in the pretest. Finally, students participated in a final session where they solved the same problems as they had in the pretest. Accuracy, latency, and self-report data were gathered in order to examine several questions.

Precursors of the discovery of the shortcut were of interest. Siegler and Crowley (1991), in their microgenetic study of the min strategy, found a <u>latency</u>

bump, a significantly longer latency on the problem immediately preceding the discovery of a new solution procedure. Therefore, we wanted to examine students' latency data to assess whether such a bump was also present in this study. Contrary to previous findings, Siegler and Crowley also found that discoveries occurred equally often on easy and hard problems. That is, a new strategy or shortcut did not tend to be used initially on more difficult problems, or impasse problems, an assumption of most learning models. As VanLehn stated: "Learning occurs only when an impasse occurs. If there is no impasse, there is no learning" (pp. 31-32, cited in Siegler & Crowley, 1991).

However, impasses or difficult problems do seem to have a role in the generalization of strategy use, according to Siegler and Crowley (1991). They found that discovery of a new strategy did not mean that the strategy would quickly become generalized to all types of problems. When given impasse problems, the students who had already discovered the min strategy quickly generalized, but the problems had no impact on the students who had not yet discovered the strategy.

Age differences in the discovery and generalization of a shortcut were also of interest. Different precursors might be required for students of different ages to discover or generalize the inversion-based shortcut, a result that might reflect an age difference in the relation between conceptual and procedural knowledge. Although the pretest revealed that there were no age differences in shortcut use, age differences in the reasons for shortcut use were hypothesized. Students in Grade 1 had great difficulties solving the large problems, and to a certain extent, the small problems that had a sum of the first two terms that was greater than 10. Students in Grade 5, however, had little difficulty solving any of the problems. Based on the age differences in ease of solving three-term problems and the lack of age differences in spontaneous use of shortcuts, necessity was hypothesized as the reason for shortcut use by the Grade 1 students who had enough conceptual knowledge about the underlying principles of arithmetic, whereas the Grade 5 students used the shortcut because it was an easier and faster solution procedure than a left-to-right calculation. The microgenetic method would help investigate such age differences by permittting examination of whether there were age-related differences in the discovery and generalization of the inversion-based shortcut. These differences would yield important information about cognitive development and the related change mechanisms. As well, these differences may have instructional implications in that types of instruction in solution procedures may differ according to age.

Students in both the experimental and control groups were given an evaluation-of-procedures task during the third session, and once again in the posttest. In this task, students were given lists of three-term problems of five types (small simple inversion, large simple inversion, small reversed inversion, large reversed inversion, and a final list that included small and large standard problems). Students were given a description of the inversion-based shortcut and were asked (a) to evaluate whether it would be appropriate for solving the problems on each list, and (b) to justify their response. This task was given for several reasons; (a) to assess whether students could recognize and justify the proper use of the shortcut independent of whether they actually used the shortcut when solving problems (b) to evaluate whether the students had conceptual knowledge about the principle of inversion, that is, whether students understand the underlying concepts about the principle, (c), to examine the relation between performance on the evaluation-of-procedures task and the use of the inversion-based shortcut, and (d) to assess whether the evaluation-of-procedures task would promote discovery or generalization of the inversion-based shortcut.

Method

Subjects

Subjects in each grade who never reported using the inversion-based shortcut in Study 1 were randomly assigned to either an experimental group or control group for Study 2. An experimental group at each age consisted of six students, half male and half female. Control groups consisted of 5 students (2 males and 3 females) in Grade 1 and 6 students (4 males and 2 females) in Grade 5. Students were randomly placed in either the experimental or control groups.

Materials and Procedure

The experimental and control groups participated in slightly different conditions, although the pretest and the posttest were identical for both groups (see Table 12).

Insert Table 12 about here

Experimental group. Five stimulus lists were used for the five sessions (see Appendix E). The same types and sizes of problems were used in the sessions as in the pretest but the problems were different from the pretest problems. The same problem appeared a maximum of two times throughout the sessions. Each stimulus list consisted of twelve inversion problems and four standard problems. Three small and 3 large problems for both simple and reversed inversion problems were included, as well as 2 small and 2 large standard problems.

The 16 problems were ordered unsystematically with the constraints that no two problems of the same type and size were presented consecutively, no more than two inversion problems of either type or two standard problems were presented consecutively, and a small inversion problem (of either type) was always presented first and last. Two list orders were used, one the reverse of the other. All problems were presented on 8 1/2" X 11" white paper with a transparent plastic cover, with a maximum of six problems per page.

In the first session, subjects were asked to solve the 16 problems and to immediately explain how they solved each problem after they had given the answer, just as they had in the pretest.

In the second session, subjects were given the second list of 16 problems. To assess whether the students were aware of the shortcut but simply did not use it, subjects were asked if there were any other ways that they could solve the inversion problems.

In the third session, the subjects were given the third set of 16 problems. At the end of the session, the subjects were given an evaluation-of-procedures task to assess whether they could recognize and justify appropriate use of the shortcut based on the inversion principle. Subjects were given, in counterbalanced order, four different lists of inversion problems: small simple inversion, large simple inversion, small reversed inversion, and large reversed inversion problems. A fifth list consisted of small and large standard problems (Appendix F). For each list, students were given a description of the inversion-based shortcut and were asked to evaluate whether it would be appropriate for solving the problems on that list, and to justify their responses (Appendix G). Subsequent to the completion of this task, subjects were given six additional problems (one of each problem size and type combination) to assess whether the evaluation-of-procedures task promoted the use of the inversion-based shortcut (Appendix H).

In the fourth session, subjects were given the fourth set of 16 problems. For those subjects who had yet to use the inversion shortcut, or were not using it regularly, direct instruction was provided (Appendix I). Direct instruction consisted of showing the subjects a list of inversion problems and telling them that b - b = 0 and therefore the answer would simply be a, what was left over (Appendix J). After direct instruction, six additional problems (one of each problem size and type combination) were given to see whether the instruction promoted shortcut use (Appendix K)

In the fifth session, subjects received the fifth set of 16 problems. For subjects who still did not use the inversion-based shortcut or who used it sporadically, a second direct instruction session was given (Appendix L). This direct instruction was similar to the first but consisted of using concrete objects (pennies) to demonstrate that b - b = 0 and only a objects were left (Appendix M). Six additional problems (one of each problem size and type combination) were subsequently given (Appendix N).

In the posttest, one week after the fifth session, subjects were asked to solve the same set of problems as they had solved in the pretest, Study 1. Following the problem-solving task, subjects were asked to do the same evaluation-of-procedures task that they had done in the third session. Subjects were videotaped during all of the five sessions and the posttest.

Control Group. The control group participated in a session at the same time as the experimental group's third session. The control group was formed to assess whether there were any schooling effects. Students may have increased their use of the inversion-based shortcut independent of the microgenetic study. Instead of solving 16 three-term problems, the control group solved 16 two-term addition and subtraction problems using the same as and bs as in the corresponding three-term problems that the experimental group was given (Appendix O). This was done in order to ensure that the control group had a similar practice period preceding the evaluation-of-procedures task. Two list orders were used, one the reverse of the other. All problems were presented on 8 1/2" X 11" white paper with a transparent plastic cover, with a maximum of six problems per page. The control group was also given the evaluation-of-procedures task subsequent to the problem-solving task.

After the problem-solving task, the control group participated in the same evaluation-of-procedures task as the experimental group. This was done in order to assess whether there were schooling effects affecting the performance on this task. The control group was given the same six additional problems after the second task that the experimental group was given. The control group also completed the posttest at the end of the study. Subjects were videotaped during the third session and the posttest.

Results and Discussion

To assess whether students had changed between the beginning and end of the study we compared accuracy, latency, and self-report data from the pretest and the posttest. If students discovered the inversion-based shortcut during the study, then accuracy data should improve, latencies should decrease, and there should be more self-reports of the use of the shortcut. Questions of interest included (a) whether the use of the inversion-based shortcut increased between the pretest and the posttest, (b) when the shortcut was discovered, (c) whether the shortcut generalized, and (d) whether there were any age differences in discovery and generalization. As well, the relation between the use of a shortcut and the ability to recognize and justify its appropriate use was of interest.

Pretest versus Posttest

Self-Reports. For the self-report data, statistical tests were not done due to floor effects on the pretest, ceiling effects in the posttest, and small sample sizes that would limit the value of these tests. According to self-reports, the use of the inversion-based shortcut increased for all groups between the pretest and the posttest. However, the experimental groups in both grades used the shortcut more frequently on the posttest than the control groups. Age differences were found for both the control and experimental groups. Overall, Grade 5 students used the shortcut more frequently than Grade 1 students in both groups (see Table 13). Grade 1 students in the experimental group used the inversion-based shortcut much more frequently on large than small problems (see Table 14). This finding is consistent with the hypothesis that younger students use the shortcut out of necessity, on large problems, and on small problems, the left-to-right solution procedure is still adequate. Grade 5 students, however, appeared to use the shortcut equally often on both sizes of problems. However, because use of the shortcut was at ceiling, there was no room for Grade 5 students in the experimental group to show an increase from small to large problems. This finding is consistent with the hypothesis that, unlike the Grade 1 students who are using the shortcut out of necessity, the Grade 5 students use the shortcut because it is simpler and faster than using a left-to-right solution procedure. In the control group, the Grade 1 students used the shortcut somewhat more often on large than small problems, consistent with the experimental group. These differences, however, were very small. They also used the shortcut more on simple inversion than reversed inversion problems. This difference in shortcut use may be because the b terms were adjacent in the former type of problem, making the possibility of using the shortcut more obvious. Grade 5 students in the control group used the shortcut almost equally often on small and large problems for the reversed inversion problems but tended to use the shortcut somewhat more often on large simple inversion problems than on small problems of that type.

Insert Table 13 about here

Insert Table 14 about here

Accuracy. As discussed in Study 1, a problem-size effect is expected if students use a left-to-right solution procedure to solve three-term problems. That is, longer latencies and lower levels of accuracy should exist for large problems as compared to small problems. If, however, the inversion-based shortcut is being used, there should be no problem-size effect because the size of the numbers in the problem are irrelevant to shortcut use. Accuracy and latency data were examined in terms of problem-size effect for experimental and control subjects in both grades.

At the pretest, none of the subjects in either group or grade reported using the shortcut. Therefore, a problem-size effect for accuracy on all problem types would be predicted. Simple contrasts were performed on all types of problems at both tests for both groups in each grade. A Bonferroni correction was used such that $\underline{p}$ values had to be less than .0004 ($\underline{p} < .05$ for 12 comparisons) for the effect to be significant. In Grade 1, at the pretest, problem-size effects were significant for each problem type within the experimental group (see Table 15). For the control group, there was only a significant problem size effect for both types of inversion problems although the problem-size effect was in the right direction for the standard problems. Although subjects were randomly placed in either the experimental or control group, the experimental group, unlike the control group, was completely inaccurate on large problems. Therefore, conclusions about the experimental and control groups must take into account the discrepancies between both groups for pretest accuracy on large problems. Proportions of accurate responses were very low for both groups. For the posttest, the experimental group showed a problem-size effect for simple inversion and standard problems. However, accuracy on simple inversion problems was higher on large than small problems, consistent with the higher reported use of the shortcut on large than small problems. For standard problems, students were unable to accurately solve any of the large problems and had much lower proportion of accurate responses on the small problems as compared to small inversion problems. Therefore, results are consistent with the use of a shortcut on the inversion problems by the posttest because students were still unable to solve standard problems for which the inversion-based shortcut was inappropriate but were able to solve inversion problems that had terms of comparable magnitude. The control group, however, showed a similar problem-size effect for all types of problems. These effects are consistent with the low use of shortcuts that they had reported. The much higher level of accuracy on small simple inversion problems is consistent with the 20 percent increase in shortcut use by the end of the study. Accuracy data for Grade 1 experimental and control groups are consistent with the experimental group's large reported increase in the use of the inversion-based shortcut as compared to the small increase reported by the control group.

Insert Table 15 about here

Grade 5 students, as in Study 1, showed very high levels of accuracy and so problem-size effects were quite small for all problem types at both pretest and posttest (see Table 16). Therefore, although tests of simple effects were computed, they must be interpreted cautiously due to ceiling effects. For the experimental group, small problems tended to be solved somewhat more accurately than large problems on both pretest and posttest with the exception of standard problems on the posttest, where large problems were solved more accurately than large problems. No explanation of this lack of problem-size effect is obvious. Although this finding would be consistent with the use of the subtraction-first shortcut, students never reported using the subtraction-first shortcut. The control group solved small problems somewhat more accurately than large problems on all problem types at both pretest and posttest except for simple inversion problems that were perfectly accurate, regardless of problem size. Because the students in both groups were so close to ceiling, all differences must be interpreted cautiously.

Insert Table 16 about here

Latency. Median latencies were also expected to have differing problem-size effects on the pretest and the posttest if the inversion-based shortcut was being used by the end of the study. Means for Grade 1 students were very high at pretest, regardless of problem type or size, although latencies for small problems were slightly faster than for large problems (see Table 17). Recall that when students could not solve the problem by the 40 second limit, a discontinuation protocol was implemented and latency was scored as 40 seconds. Both the experimental and control groups showed a significant problem-size effect on reversed inversion problems, in that small problems were solved more quickly than large problems. Because latencies were so close to ceiling, little can be concluded from these results. In the posttest, the experimental group showed no significant problem-size effects for either type of inversion problems despite absence of ceiling effects. A large problem-size effect was found on standard problems. These results are consistent with the experimental group's self-reports of an increase in the use of the inversion-based shortcut. The control group had fairly large problem-size effects on all problem types yet the effect was smaller on simple inversion problems, which is consistent with the slightly higher reported use of the shortcut on simple inversion problems.

Insert Table 17 about here

The Grade 5 experimental group showed no significant problem-size effects on the inversion problems on the pretest, but the effects went in different directions (see Table 18). On simple inversion problems, latencies were slightly faster on small than large problems whereas the inverse tendency was found for the reversed inversion problems. A problem-size effect was found on the standard problems, with faster latencies on small than large problems. The control group had significant problem-size effects for both simple inversion problems and standard problems. Although not significant, the problem-size effect for the reversed inversion problems was in the same direction. The problem-size effects for the standard problems parallel the effects found on the accuracy data for both groups. However, the other latency results are inconsistent with the accuracy data.

Insert Table 18 about here

In the posttest, the problem-size effects were very small and insignificant on all three types of problems for the experimental group. The extremely small problem-size effect on the large standard problems is puzzling and no explanation is obvious. This result parallels the finding that accuracy on standards was significantly higher on large than small problems. The control group, however, showed negligible problem-size effects on the inversion problems and a larger effect on the standard problems, a pattern consistent with the use of a shortcut on inversion problems.

Comparisons of means for accuracy, latency, and self-report data between the pretest and the posttest indicate that at least some of the students of both grades in the experimental group were using the inversion-based shortcut during the posttest. The data also indicate that there was also some use of the shortcut for the control groups in each grade, although levels of use were much higher for the experimental groups. However, these data give little information about how the students made the change between no shortcut use at the pretest and at least some shortcut use by the posttest. Examining each student's shortcut use on a session-by-session basis helps to reveal precursors of change.

Session-by-Session Analyses

For each problem type and size combination, examinations focused on the session in which each student (a) first used the shortcut (discovery), (b) first used the shortcut on at least 75% of the inversion problems, including at least once on each problem type and size combination (frequent use), and (c) first started using the shortcut frequently and used it frequently on all subsequent sessions, (consistent use) (see Tables 19 and 20). For example, Subject 545 discovered the shortcut on all problem type and size combinations during session 2 except for small simple inversion problems, for which the shortcut was discovered during session 3. Subject 545 became a frequent user on all problem type and size combinations during session 3. The shortcut was used consistently for all types on problems beginning in session 3 except for small simple inversion problems, which were not solved consistently with the shortcut until session 5.

Insert Table 19 about here

Insert Table 20 about here

All of the experimental students discovered the inversion-based shortcut during the microgenetic study. However, an age difference may exist in the precursors of shortcut discovery. Two Grade 1 students (101 and 125) and three Grade 5 students (506, 513, and 545) discovered the shortcut in either the first or the second sessions. The remaining Grade 5 students discovered the shortcut after the evaluation-of-procedures task, which involved demonstrations of the inversion-based shortcut and asking the subjects whether the shortcut would work to solve inversion problems. Thus, this task was sufficient for the Grade 5 students to discover the shortcut, but it had no apparent effect on the Grade 1 students who had yet to discover the shortcut. For these students, direct instruction was necessary for shortcut discovery. In Sessions 4 and 5, the students were shown how to solve problems using the inversion-based shortcut. For example, Subject 105 did not discover the shortcut until the problem-solving task immediately following the first direct instruction task. Therefore, for students who do not spontaneously use the inversion-based shortcut, different conditions for discovery seem to be necessary for students of different ages. As well, individual differences exist within each grade.

Another question of interest for the discovery of a shortcut is whether a certain problem type or size was conducive to shortcut discovery. Overall, for problem type, the use of the shortcut on both problem types was most often discovered during the same session for each problem size, e.g., the discovery of the shortcut on large simple inversion problems usually occurred during the same session as the discovery on large reversed problems. For example, Subjects 123 and 530 discovered the shortcut on both types during the same session. For problem size, the inversion-based shortcut tended to be discovered for both problem sizes in the same session although some subjects (4 out of 12) discovered the shortcut on large before small problems. For example, Subjects 102 and 506 discovered the shortcut on large problems of both types during the first session but only discovered the shortcut on small problems during subsequent sessions. Problem type seems to be an irrelevant variable in the discovery of a shortcut because the shortcut tends to be discovered on both problem types during the same session. However, problem size may have an effect on discovery because some of the students tended to discover the inversion-based shortcut on large before small problems. This finding is consistent with the use of a shortcut on problems where it is more effective or useful than the left-to-right solution procedure.

After students discovered the inversion-based shortcut, they did not necessarily use the shortcut regularly on subsequent inversion problems. Students were classified as <u>frequent users</u> if they used the shortcut on at least 8 of the 12 inversion problems during a session, and at least once on each problem type and size combination. Students were classified as <u>consistent users</u> when they used the shortcut frequently on all sessions, starting on the first session that they were classified as frequent users and remained frequent users in all subsequent sessions. Because some Grade 1 students did not start using the shortcut frequently until the posttest (e.g., Subject 138), no consistent use classification could be made. For each problem type and size combination, the session in which the use of the shortcut became frequent and/or consistent was noted.

Most students became consistent users at the same time as they became frequent users (e.g., Subjects 125 and 549). Frequent and consistent use appeared to follow the evaluation-of-procedures task for Grade 5 students (4 out of 6). For example, Subject 530 discovered the shortcut immediately following the evaluation-of-procedures task and became both a frequent and consistent user on the subsequent session. Grade 1 students, however, did not always become frequent users, let alone consistent users, especially on small problems. For example, Subject 121 never became a frequent or consistent user on small problems of both types, although he did for large problems. This size difference in the generalization of the use of a shortcut is consistent with the hypothesis proposed in Study 1 that at least some Grade 1 students who had enough conceptual knowledge about the underlying principles of arithmetic use the shortcut out of necessity. Grade 1 students had great difficulties solving large inversion problems if the shortcut was not used. Exposure to the problem-solving sessions may have been enough to promote conceptual understanding of the principle of inversion. Grade 5 students, however, could solve problems of both sizes equally well and probably used the shortcut because it was faster and easier than a left-to-right solution procedure, and not out of necessity, and hence the infrequent size differences in generalization.

In their study, Siegler and Crowley (1991) found that there was a significant increase in latency just before the discovery of the min strategy, an increase that they called a latency bump. Latency data were examined to see whether there was such a bump. First, mean latencies for all subjects grouped together were examined and then the latencies for each subject were examined. Finally, latencies for each subject were separated by problem type and size and were examined. No bump in latency was found in these data. However, failure to find such a bump should not be taken as strong evidence against the conclusions of Siegler and Crowley. In this study, six different problem type and size combinations were used, whereas in Siegler and Crowley's study only one type of problem was used. All the combinations were distributed within each session such that the same combination never appeared consecutively. Therefore, to compare the latency of the problem preceding the discovery of the shortcut use would be comparing two different problem type and size combinations. As well, students sometimes discovered the shortcut on the first problem in a session. Latencies were usually slowest at the beginning, regardless of problem type or size, and fastest at the end. Therefore, to compare the latency of the problem preceding the discovery to the latency of the first problem in the session would have been inappropriate.

Siegler and Crowley (1991) also found that no incorrect strategies were ever used. In this study, however, some of the Grade 1 students (Subjects 125 and 138) overgeneralized the shortcut by using it, incorrectly, on standard problems. For example, these students would say the answer was 3 for the problem 3 + 25 - 21 because "25 take away 21 leaves 3". Subject 125 overgeneralized on all sessions subsequent to, and including, session 2 while subject 138 only overgeneralized on sessions 4 and 5. No students in Grade 5 ever misused the shortcut on standard problems. Therefore, the overgeneralization of younger students is consistent with the notion that there age differences in students who do not spontaneously use the inversion-based shortcut. The difficulty of solving the large standard problems may have led younger students to use the shortcut on standard problems as a last resort. Individual differences in how strongly the students felt that they had to answer the problem even if they did not know how to may have caused some students to use incorrect solution procedures to solve the problems, whereas other students simply may have stated that they could not solve the problems but if they could, they would use a left-to-right solution procedure.

In Study 1, a positive correlation between the inversion-based shortcut and the subtraction-first shortcut was found. All of the subjects in the microgenetic study discovered the inversion-based shortcut, therefore it would be possible that the subtraction-first shortcut also would be discovered. The twelve experimental group subjects in this microgenetic study solved a total of 128 three-term problems each. However, the subtraction-first shortcut was used only once, and at that incorrectly, by a Grade 1 student. There was no use of the subtraction-first shortcut even after the use of the inversion-based shortcut had generalized. Therefore, subjects are able to acquire the inversion-based shortcut either through practice, exposure to the evaluation-of-procedures task, or direct instruction, but none of these conditions immediately led to the use of a similar shortcut, the subtraction-first shortcut. This result is consistent with the finding in Study 1 that the use of the inversion-based shortcut was much higher than that of the subtraction-first shortcut.

Evaluation-of-Procedures Task

Comparisons were done on performance on the evaluation-of-procedures task between session 3 and the posttest. Students in both the experimental and the control groups were asked to do this task immediately after the problem-solving task in session 3 and the posttest. On the recognition component, students were classified as recognizers if they said that the shortcut would work on at least 3 of the 4 inversion problem type and size combinations and also said that the shortcut would not work on the standard problems. Students who said the shortcut would work on all 5 lists of problems were classified as false positives. All students who did not fit in these two categories were classified as other. This category included students who were unable to say whether the shortcut would work or not on more than one list or gave ambiguous responses on more than one list.

On the justification component of the evaluation of procedures task, students were asked to justify the answer that they had just given on the recognition component. Students who were able to justify their answers were classified as <u>justifiers</u>, students who said the shortcut would work on all 5 lists of problems and gave an explanation based on the principle of inversion for all 5 lists and were classified as <u>false positives</u>, students who gave an appropriate justification on the 4 inversion problem lists but were unable to provide a justification for their answer to the standard problem list were classified as <u>incomplete justifiers</u>, and students who did not fit into either category were classified as <u>other</u>. Students who were classified as other included students who were unable to justify on more than one list or gave more than one ambiguous response.

During session 3, performance on the recognition component of the evaluation-of-procedures task was similar for both the experimental and control groups in Grade 1 (see Tables 21 and 22). Only one student in each group was a recognizer and most of the others in the experimental group and in the control group were false positives. By the posttest, performance on the recognition task had improved slightly for the Grade 1 experimental group: Half of the students were now recognizers. The control group showed no changes. Students in the experimental group who overgeneralized the shortcut on the problem-solving task were equally divided between recognizers and false-positives.

Students in Grade 1 tended to agree that the use of a shortcut was appropriate even on standard problems. It may be that students were simply agreeing with everything that the experimenter said, rather than actually believing that the shortcut was appropriate for solving standard problems. Because students were not given a practice session before this task to evaluate whether they would always agree with the experimenter, caution may be warranted for the interpretation of these findings.

Insert Table 21 about here

Insert Table 22 about here

Students in Grade 1 had even more difficulty with the justification task than the recognition task. Three students in each of the experimental and control groups were unable to give justifications on more than one list on the pretest. The remaining students in the experimental group and one student in the control group justified were false positives. One control group student was classified as an incomplete justifier. By the posttest, two students in each group were justifiers. Half of the experimental group and one of the students in the control group were false positives. One student in the experimental group and two in the control group were still unable to provide justifications.

For Grade 5, on the evaluation task in session 3, all of the students in the experimental group and all but one of the students in the control group were recognizers (see Tables 23 and 24). By the posttest, the pattern was reversed: All of the students in the control group and all but one of the students in the experimental group were recognizers. Overall, performance of the Grade 5 students was at ceiling. Therefore, by Grade 5, students have mastered the ability to appropriately evaluate and justify the use of a shortcut.

Insert Table 23 about here

Insert Table 24 about here

In Grade 5, in the justification part of the evaluation-of-procedures task in session 3, all of the experimental group and half of the control group were justifiers. By the posttest, four students in the experimental group and five students in the control group were justifiers. Again, ceiling effects were found for the older students.

The microgenetic sessions appear to have had little effect on performance for the evaluation-of-procedures task for either Grade 1 or Grade 5 students. Differences between the experimental and control groups were minimal. However, there were large differences between age groups. Grade 1 students had a fairly high level of difficulty for both components of the evaluation-of-procedures task. Although Grade 1 students in the experimental group did improve, to a certain extent, between session 3 and the posttest, so did the control group. Grade 5 students were at ceiling, regardless of whether they were in the experimental or control group.

Although the Grade 1 and 5 students were identical in that none of them spontaneously used the inversion-based shortcut on the pretest, the findings from the evaluation-of-procedures task are consistent with the hypothesis that these Grade 1 and 5 students were not alike. Grade 1 students had great difficulties with the evaluation-of-procedures task whereas Grade 5 students performed well. Also, even though some of the students in Grade 1 discovered the inversion-based shortcut early in the microgenetic study (subjects 102 and 125), they were both classified as false-positives on the recognition task as well as the justification task. Therefore, there appears to be little relation between the discovery of a shortcut and performance on the evaluation-of-procedures task for the Grade 1 students. Although the Grade 1 students were all using the shortcut, at least to some extent, by the posttest, they could not properly recognize and justify the appropriate use of the shortcut.

For Grade 5, performance on the evaluation-of-procedures task was at ceiling. However, for those students who had yet to discover the inversion-based shortcut, exposure to the evaluation-of-procedures task was sufficient to promote shortcut discovery in subsequent problem-solving. Therefore, the evaluation-of-procedures task influenced performance on the problem-solving task for those older students who had not previously discovered the shortcut.

Based on these age differences, the relation between the ability to use a shortcut based on the principle of inversion and the ability to recognize and justify the appropriate use of the shortcut may be complex. At a young age, the ability to recognize and justify the shortcut is unnecessary for using the shortcut. As students grow older, the ability to recognize and justify appropriate shortcut use is mastered. However,if they have not yet used the shortcut, the ability to recognize and justify the shortcut may be necessary to promote shortcut use.

Conclusions

Based on these results, there seems to be evidence suggesting that the discovery and generalization of a shortcut occur under different circumstances for children of different ages. These findings have important implications for our knowledge about the mechanisms of change as well as the instruction of solution procedures to students in different grades. Demonstrating a solution procedure and asking students to evaluate and justify its appropriateness seems to be a sufficient change mechanism and instructional method for teaching a new solution procedure to older students. However, younger students need direct instruction of the shortcut as a change mechanism in order to start using the shortcut. As well, younger students normally have little exposure to difficult problems in arithmetic and yet difficult problems often yielded "impasse-driven learning" in both discovery and generalization of a shortcut. Therefore, exposing younger students to more difficult problems may have positive effects in the acquisition of new solution procedures. However, these results also demonstrate the individual variability in the precursors of the discovery and generalization of the inversion-based shortcut within each age.

Although students in Grade 1 were all capable of using the inversion-based shortcut on the problem-solving task, performance was not very good on the evaluation-of-procedures task. Therefore it is questionable whether these students actually "learned" the principle of inversion. Rather, they were able to implement a solution procedure based on the principle of inversion without always being able to understand the principle.

For students who do not spontaneously use the inversion-based shortcut, the relation between use of a shortcut and recognition and justification of that shortcut appears to differ according to the age of the students. Use of the inversion-based shortcut appears to precede the ability to recognize and justify the appropriate use of the shortcut for most Grade 1 students, whereas recognition and justification appears to precede use of the shortcut for the Grade 5 students. These age differences may indicate a complex relation between the development of the ability to use a shortcut based on the principle of inversion and the ability to recognize and justify the appropriate use of the shortcut. At a young age, the ability to recognize and justify the shortcut is unnecessary for using the shortcut. As students grow older, the ability to recognize and justify appropriate shortcut use is mastered. However,if they have not yet used the shortcut, the ability to recognize and justify the shortcut may be necessary to promote shortcut use. Again, this differing relation between grades may be due to differing reasons for using the shortcut. Grade 1 students need to find a way to solve these problems and they do not need to understand the principle of inversion fully before using the shortcut. Grade 5 students do not need to find a novel solution procedure to solve the problems and therefore may not use the shortcut until they are exposed to the evaluation-of-procedures task.

Although the findings of this study are compelling, because of the small sample sizes and the floor and ceiling effects that were found in accuracy, latency, and self-report data as well as in the evaluation-of-procedures task, these results should be interpreted carefully. Statistical analyses were, for the most part, inappropriate because of low sample size and low variability. This study needs to be replicated using larger sample sizes. Also, ceiling effects that were found consistently in Grade 5 need to be removed by including more difficult problems in the problem-solving task, e.g., problems such as 32 + 129 - 129. Also, not all of the younger students became frequent or consistent users by the end of the study. A microgenetic study should be long enough to encompass "the beginning of the change to the time at which it reaches a relatively stable state" (Siegler & Crowley, 1991). Finally, we need to conclusively establish that younger students actually believe that the inversion-based shortcut is appropriate for solving standard problems on the evaluation-of-procedures task and not that they are simply agreeing with everything that the experimenter says.

Study 3

The posttest was given in order to evaluate whether there were changes in the use of the inversion-based shortcut between the beginning and the end of the experiment for the experimental and control groups of Study 2. For those students who used the shortcut frequently or infrequently in Study 1, Study 3 was designed to assess the stability of shortcut use over time. Once the students start using the shortcut, they should continue using it because the shortcut is easier and faster than a left-to-right solution procedure. In Siegler's model of strategy choice, once students start retrieving answers to two-term problems (the most efficient way to solve the problems), they continue to use retrieval as their solution procedure unless they are uncertain, and then they use backup procedures (Siegler & Shipley, 1992; Siegler & Shrager, 1984). For the three-term inversion problems utilized in this study, the inversion-based shortcut is the most efficient solution procedure. Once students have used the shortcut, their use of the shortcut should remain stable or increase over time, especially on large problems. As well, results from Study 2 indicate that once students start using the shortcut, they continue to do so.

Whether use of the shortcut is stable or even increases over time is of interest for several reasons.

Development is generally thought of as progressing through stages, and that like riding a bicycle, once a child learns how, he never unlearns how, or regresses to a previous stage. Therefore, the stability or increase of use of the shortcut lends itself well to a theory of stages in development. An educational implication is that stability of the use of the shortcut indicates that once a child learns a new solution procedure, no reinforcement by the teacher is needed for continued use of the shortcut. Finally, although analyses of cross-sectional data could show an increase or equal use of the shortcut over time, these analyses yield little information about individual differences. How each student changes or does not change is of interest. Therefore, whether Users and Variable Users retained their classification or whether they became Users, Variable Users, or Non-users was also examined.

The evaluation-of-procedures task used in Study 2 was employed again in Study 3 to assess whether students in both grades were capable of recognizing and justifying the proper use of the inversion-based shortcut. This task was designed to assess a different type of understanding than the problem-solving task. Although a student may not demonstrate understanding of the inversion principle by using the associated shortcut, that student may still be able to understand the principle in that he or she can recognize and justify its appropriate use.

All but one of the subjects who had been labelled Users or Variable Users in the pretest were included in the pretest and posttest comparison. The only exception was a Grade 1 female who had been inadvertently placed in the control group of Study 2 despite having been categorized as a User in Study 1. The subject's data was not included in Study 2 either.

It was expected that the use of the inversion-based shortcut would remain fairly stable between the pretest and the posttest, although an increase of shortcut use could be possible due to generalization for the Variable Users to different problem type and size combinations. Moreover, it was expected that latency and accuracy data might improve between the pretest and the posttest due to increased use of the inversion-based and subtraction-first shortcuts, or alternatively because of practice effects.

Because all of the students involved in the analyses were capable of using the inversion-based shortcut, it was expected that the students would also be successful at the evaluation-of-procedures task.

However, based on the results of Study 2, students in Grade 1 had difficulty with this task, even if they had already discovered the inversion-based shortcut. Although the younger students used the shortcut in the microgenetic study, it appeared that the ability to use the inversion-based shortcut was independent of the ability to recognize and justify its appropriate use. Another question of interest was whether the use of the subtraction-first shortcut would increase between pretest and posttest.

Method

Subjects

Subjects were all of the students who were classified as Users and Variable Users in the pretest, except for one female Grade 1 User. Therefore, 26 Grade 1 and 27 Grade 5 students participated.

Materials and Procedure

Subjects were given the same list of 24 problems to solve as in Study 1 (Appendix A). In addition, the students were given the evaluation-of-procedures task that was used in the third session of Study 2 (Appendix G). The same procedures were used as in Studies 1 and 2. Subjects were then thanked for participating in the study and told that their participation would help the experimenter learn about how students think about and solve arithmetic problems.

Results and Discussion

To assess whether the use of a shortcut was stable over time, pretest and posttest data were compared. Students' use of the inversion-based shortcut was expected to remain stable or increase only slightly due to practice, because the results of previous studies indicate that shortcut use does not decrease with age (Bisanz, et al., 1989; 1992; Dhaliwal, 1989). To assess possible changes over time, effects and interactions involving test (pretest vs. posttest) are described. Based on the conclusions of Studies 1 and 2 that self-report data are veridical, these data are examined first. Latency and accuracy data are examined later for corroboration.

Self-Reports of the Inversion-based Shortcut

Proportions of self-reported use of the inversion-based shortcut were subjected to a 2(Grade: 1 and 5) x 2(Test: pretest and posttest) x 3(Type: simple inversion, reversed inversion, and standard) x 2(Size: small and large) analysis of variance with repeated measures on the last three variables. Reported use of the shortcut was higher on the posttest than on the pretest (.61 vs. .77), $F(1,51) = 24.22$, $p < .01$. Test and size interacted, $F(1,51) = 10.32$, $p < .01$ (see

Table 25). The inversion-based shortcut was reported more frequently on the posttest than the pretest for both small and large problems (ps < .05), and students reported using the shortcut more often on large than small problems on both the pretest and the posttest (ps < .01). The differences between use of the shortcut on small and large problems was smaller at the posttest than the pretest, however.

Insert Table 25 about here

The shortcut seems to be more useful for solving large problems that are extremely difficult for younger students and where computational errors are more likely to occur for older students. The use of the shortcut was lower on reversed inversion problems than on simple inversion problems on both the pretest and the posttest (.65 vs. .57 for the pretest and .83 and .71 for the posttest), $F(1,51) = 11.91$, $p < .01$. The lack of main effect due to age or an interaction between age and test is notable. Students who spontaneously use the shortcut appear to perform similarly over time, regardless of age.

Self-Reports of the Subtraction-first Shortcut

Proportions of reported use of the subtraction-first shortcut on standard problems were subjected to a 2(Grade: 1 and 5) x 2(Test: pretest and posttest) x 2(Size: small and large) analysis of variance with repeated measures on test and size. Students reported using the shortcut more frequently on the posttest than on the pretest (.16 vs. .30), $F(1,49) = 18.08$, $p < .01$, and this effect interacted with grade, $F(1,49) = 7.10$, $p < .01$ (see Table 26). Grade 5 students reported using the shortcut more than the Grade 1 students on both the pretest and the posttest (ps < .01). As well, students in both grades reported shortcut use more on the posttest than the pretest (ps < .01), although differences between pretest and posttest were greater for the older students.

Insert Table 26 about here

Students in Grade 1 used the subtraction-first shortcut very infrequently during both the pretest and the posttest. The age difference in the use of the subtraction-first shortcut may be due to older students' understanding of the principle of associativity. That is, the younger students may not realize that solving the problem by subtracting the last term from the middle term and then adding the first term would give them the same answer as if they had successively added and subtracted.

Accuracy

Proportions of correct responses were subjected to a 2(Grade: 1 and 5) x 2(Test: pretest and posttest) x 3(Type: simple inversion, reversed inversion, and standard) x 2(Size: small and large) analysis of variance with repeated measures on the last three variables. No effect of test was found (.75 vs. .78), $F(1,51) = 1.48$, $p > .05$, indicating that overall accuracy remained relatively stable over time. However, an interaction between grade and test was found, $F(1,51) = 4.40$, $p < .05$. This interaction was included in the three-way interaction between grade, test, and size, $F(1,51) = 5.29$, $p < .05$. The problem-size effect decreased from pretest to posttest for Grade 1 students but not for older students, who were near ceiling on both tests (see Table 27). The improvement in accuracy for the younger students, especially on large problems, is consistent with the notion that use of the inversion-based shortcut increased, especially on large problems, because the Grade 1 students had great difficulty solving the large standard problems on both the pretest and the posttest (mean proportion accuracies were .08 and .17 on the pretest and posttest, respectively). Alternatively, Grade 1 students' problem-solving abilities may have simply improved over time. Grade 5 students were equally accurate on both sizes of problems on the pretest but tended to be slightly more accurate on small than large problems by the posttest. Accuracy was equal across test for small problems but accuracy on large problems tended to be slightly greater on the pretest than the posttest. Grade 5 students were extremely close to ceiling for accuracy measures on both the pretest and the posttest, so their results should be interpreted cautiously.

Insert Table 27 about here

Latency

Separate latency analyses were performed for Grades 1 and 5. Medians of the correct response latencies for each grade were subjected to a 2(Sex: male and female) x 2(Test: pretest and posttest) x 3(Type: simple inversion, reversed inversion, and standard) x 2(Size: small and large) analysis of variance with repeated measures on the last three variables. In Grade 1, the students solved problems more quickly in the posttest than the pretest (21.7 s vs. 18.4 s), $F(1,21) = 5.04$, $p < .05$). No interactions involving test were found.

If the Grade 1 students were simply improving in their ability to add and subtract, there should have also been an improvement on the large standard problems. However, latencies on large standard problems were near ceiling not only on the pretest but also on the posttest (38.0 s and 34.8 s). Therefore, the improvement in solution latency is consistent with the notion that use of the inversion-based shortcut remained stable and may even have increased.

The overall decrease in latency in Grade 5 indicates that students are improving over time (4.4 s vs. 2.9 s), $F(1,25) = 41.13$, $p < .01$. This main effect is qualified by interactions between test and size, $F(1,25) = 8.59$, $p < .01$, and among test, type, and size, $F(2,50) = 7.77$, $p < .01$ (see Table 28). Students were faster on all problem type and size combinations on the posttest than the pretest ($ps < .01$). The shorter solution latencies for large problems than small problems for simple inversion problems ($ps < .01$) is consistent with the higher use of the shortcut on large than small problems in both the pretest and the posttest. The lack of problem-size effect on the reversed inversion problems in the posttest is also consistent with the increased use of the inversion-based shortcut on large problems. As well, the lack of problem-size effect on standard problems is consistent with the increased use of the subtraction-first shortcut in the posttest, especially on large problems. This finding parallels the lack of problem-size effect for accuracy data. Finally, the results are consistent with the notion that the use of a shortcut is greater on both types of inversion problems than on standard problems and that the inversion-based shortcut is more frequent on simple inversion than on reversed inversion problems.

Insert Table 28 about here

Individual Differences

As in the pretest, students were classified as Users if they used the shortcut on more than 75% of the 12 inversion problems and they used the shortcut at least once on each inversion problem type and size combination. Variable Users were students who used the shortcut at least once but less than 75% of the time. Non-users were students who never used the shortcut. On the pretest, no age differences were found; an almost equal number of students in each grade were classified as Users, Variable-Users, and Non-users. Classifications were compared on the pretest and posttest for the Users and Variable Users (see Table 29). Although there still appear to be no age differences in the User group, Variable Users in Grade 5 tended to become Users whereas the Variable Users in Grade 1 were almost equally likely to remain Variable Users or to become Users or Non-Users.

Insert Table 29 about here

These data indicate that Users in both age groups are relatively stable. However, shortcut use does not appear to be stable for Variable Users in both grades, and this instability is different across age. Grade 1 Variable Users are equally capable of becoming Non-users or Users while Grade 5 Variable Users either remain the same or increase their usage of the shortcut. Therefore, it appears that, as for the Non-users in Study 2, there are age differences in the Variable Users in Study 3.

<u>Evaluation-of-Procedures Task</u>

Several age differences were found in the evaluation-of-procedures task. On the recognition component of the task, students were classified as recognizers, that is, being able to recognize appropriate shortcut use, if they agreed that the shortcut would work on at least 3 of the 4 lists of inversion problems and did not agree that the shortcut would work on the list of standard problems. Students were coded as false positives if, in addition to agreeing at least 3 out of 4 times that the shortcut would work on the inversion problems, they also agreed incorrectly that using the inversion-based shortcut would be appropriate for solving the list of standard problems.

For the recognition component of the task, only 11 of the 28 Grade 1 students were able to recognize the appropriate use of the inversion-based shortcut, whereas all of the 28 Grade 5 students were able to recognize appropriate shortcut use. Moreover, 16 of the 28 Grade 1 students agreed that using the shortcut would be appropriate on the standard problem list and were thus classified as false positives, whereas none of the Grade 5 students showed this pattern. The remaining Grade 1 student was classified as other because she was unable to evaluate whether the shortcut was appropriate on the list of standard problems. Based on these results, it appears that many of the younger students are unable to recognize the appropriate use of the inversion-based shortcut. Students, however, were not exposed to other situations where agreeing with the experimenter would be clearly wrong before doing the evaluation-of-procedures task. Therefore, it is necessary to establish whether younger students are simply agreeing with everything or whether they really think that the inversion-based shortcut is appropriate for solving standard problems.

Differences in performance on the recognition portion of the evaluation-of-procedures task based on the frequency of shortcut use were examined. There were no effects of user group for the Grade 5 students because all 28 students were able to evaluate the appropriate use of the shortcut. In Grade 1 however, the number of students who could recognize appropriate shortcut use or were false-positives was roughly equal for Users and Variable Users (see Table 30). There was a slight tendency for Non-users to be false positives. Therefore, using the shortcut frequently or infrequently seems to be unrelated to performance on this task.

Insert Table 30 about here

For the justification portion of the evaluation-procedures task, students were asked to justify the response that they had given to the recognition portion. Again, there were age differences in performance. In Grade 5, 27 of the 28 students were able to correctly justify their evaluations whereas only 8 of the 28 Grade 1 students were able to do so (see Table 31). These students were classified as justifiers. The remaining Grade 5 student did not use the principle of inversion to justify her answers. In Grade 1, 7 of the students used the inversion principle to explain why the shortcut would work, even on the standard problem list. These students were classified as false positives. Another 8 students used the inversion principle to explain why the shortcut would work on at least 3 of the 4 lists of inversion problems but were unable to provide any justification for their response to the standard list and were thus classified as incomplete justifiers. The 5 remaining students were unable to justify their responses on at least two of the inversion lists as well as on the standard list and were classified as other.

Insert Table 31 about here

We examined whether there were differences in performance for the justification portion of the evaluation-of-procedures task based on user groups. No differences were found in Grade 5 because all but one of the students were able to justify their responses. In Grade 1, all of the students seemed fairly equally distributed across the four categories, although the Users tended not to be classified as "other".

The justification task was very difficult for the younger children, and performance seemed to be independent of whether the students used the inversion-based shortcut frequently or infrequently on the problem-solving task. The evaluation-of-procedures task, as a whole, seems to assess understanding of the inversion principle that is independent of the implementation of the inversion principle for problem-solving for the younger students. For Grade 5 students, it appears that recognition and justification of the inversion-based shortcut are mastered more easily and probably earlier than implementation of the shortcut, as was concluded in Study 2.

Conclusions

The increase in use of the inversion-based shortcut is consistent with the generalization of the shortcut to other problem type and size combinations on the part of the Variable Users. For the Users, shortcut use appears to remain relatively stable. Therefore, once most students start using the inversion-based shortcut, they either remain stable or increase their usage of the shortcut. It appears that once students find a fast and errorless way to solve the inversion problems, they continue or increase their use of it. Once students start using the shortcut, use will increase or remain stable because students have the ability to recognize that the shortcut is the most efficient solution procedure on inversion problems. The increase in use of the subtraction-first shortcut is consistent with the generalization of the shortcut, especially on large problems. Results of the analyses on the accuracy and latency measures were compatible with the analyses of the self-reported use of the inversion-based and subtraction-first shortcuts, yielding more positive evidence for the veridicality of self-reports. The classification of Users remained relatively stable over time. Variable Users in Grade 5 tended to increase their use of the inversion-based shortcut. However, Variable Users in Grade 1 were equally likely to increase, decrease, or maintain their use of the shortcut. This finding may indicate that these younger students were less likely to generalize shortcut use and mostly used the shortcut, when necessary, on the large problems.

Based on the results of the evaluation-of-procedures task, students in Grade 5 appeared to have mastered the recognition and justification of the inversion-based shortcut. Results from Study 2 indicate that perhaps, for older students, the ability to recognize and justify the appropriate use of the shortcut may precede the use of the shortcut on problem-solving tasks. Second, most of the Grade 1 students had difficulty with the evaluation-of-procedures task, although they were able to use the shortcut on the problem-solving task. Therefore, use of the inversion principle appears to precede the ability to recognize and justify the appropriate use of the principle. However, it appears that Users tend not to be false positive. That younger students actually believe that the inversion-based shortcut is appropriate for solving standard has not yet been conclusively established. Students may simply be agreeing with the experimenter. It is clear, however, that the results of the evaluation-of-procedures task also lend credence to the hypothesis that although there are no age differences in the use of the inversion-based shortcut, age differences exist in the reasons for using the shortcut and the degree of understanding about the inversion principle that students possess.

General Discussion

Several questions were examined in this experiment. The main purposes of Study 1 were to examine age differences in students' spontaneous use of a shortcut based on the principle of inversion and the veridicality of student's self-reports. Study 2 was done in order to examine the precursors of discovery and generalization of the shortcut and the developmental relation between the ability to use the shortcut and the ability to recognize and justify its appropriate use. Finally, the purposes of Study 3 were to assess stability of shortcut use over time as well as to examine age differences in the ability to recognize and justify appropriate shortcut use.

In Study 1, students in Grades 1 and 5 were asked to solve a set of three-term problems. Accuracy, latency, and self-report data were collected. No age differences were found in the use of the inversion-based shortcut. No such result has previously been reported. Heretofore, older students have always used the shortcut more frequently than the younger students (Bisanz, et al., 1992; Dhaliwal, 1989). Although there were no age differences in Study 1, it was hypothesized that the reasons for using the shortcut were age differentiated. That is, Grade 1 students who possessed sufficient conceptual knowledge about the underlying principles of arithmetic might have been using the shortcut because otherwise they could not solve the large problems using successive addition and subtraction. The Grade 5 students, however, might have been using the shortcut because it was a faster and easier solution procedure than successive addition and subtraction. However, using the shortcut was not necessary for the Grade 5 students to solve the problems because successive addition and subtraction was an adequate solution procedure. Therefore, conceptual knowledge may also have differred in Grade 5 but alternatively, students may have been so proficient at successive addition and subtraction that the shortcut was not deemed to be a better solution procedure.

Finally, the analyses revealed that patterns across accuracy, latency, and self-report data were similar. Self-reports appear to be, at the very least, reasonably veridical. This finding is contrary to Cooney and Ladd (1992) and Russo, Johnson, and Stephens (1989), who concluded that the use of self-reports as data is questionable in research on mental arithmetic.

Despite the cautions of the aforementioned authors, self-report data, at least in this study, appear to be good measures of shortcut use. As well, the self-report data revealed that another shortcut was being used, the subtraction-first shortcut, to solve standard problems. This shortcut has not yet been reported and would never have been found without the self-report data.

In Study 2, students who did not use the inversion-based shortcut during Study 1 were included in a microgenetic study. Microgenetic studies are characterized by intensive study of a change while it is occuring. These studies can yield important information about the precursors and mechanisms of change. Use of the inversion-based shortcut was compared between pretest and posttest. By the end of this study there were age differences in the use of the shortcut. Students in Grade 5 used the shortcut more frequently than the Grade 1 students but used the shortcut equally on both problem sizes. In contrast, the younger students tended to use the shortcut much more frequently on large than small problems. Again, this finding is consistent with the hypothesis that the younger students who have the necessary conceptual knowledge are using the shortcut out of necessity while the older students use the shortcut because it is the easiest and fastest way to solve the inversion problems. Accuracy and latency data paralleled the self-report data, thereby yielding more evidence for the veridicality of verbal protocols.

Each student's performance was examined session by session to determine precursors of discovery and generalization of the shortcut occurred. Again, age differences were found. The evaluation-of-procedures task was sufficient for discovery and generalization of the shortcut for Grade 5 students. Their performance was near ceiling on both the recognition and justification components. Therefore, the ability to recognize and justify the appropriate use of the shortcut may precede shortcut use for the older students in this study.

Discovery and generalization were unaffected by exposure to the evaluation-of-procedures task for the Grade 1 students. Performance on this task was weak. The ability to use the shortcut appears to precede the ability to recognize and justify the appropriate use of the inversion-based shortcut. Direct instruction, however, was sufficient to promote discovery and generalization. Although students in both grades were equal in performance during the pretest (none of them used the shortcut), students of both ages are obviously not alike in their discovery and generalization of the inversion-based shortcut. Therefore, change mechanisms are different for Grade 1 and 5 students.

The relation between the ability to use the shortcut and the ability to recognize and justify its appropriate use appears to be complex. At a young age, shortcut use does not relate to the ability to recognize and justify its use. With age, the ability to recognize and justify appropriate use of the shortcut is mastered. By Grade 5, ability to use the shortcut and to recognize and justify apppropriate shortcut use appears to be independent for students who spontaneously used the shortcut, as was found in Study 3. For those Grade 5 students who did not spontaneously use the shortcut, the opportunity to recognize and justify appropriate shortcut use appears to promote shortcut use on subsequent problem-solving tasks.

Based on the results of Studies 1 and 2, discovery of a shortcut, at least for the younger students, seems to operate through impasse-driven learning. This is contrary to Siegler and Crowley's (1991) finding that discoveries followed successes as well as failures. Siegler and Crowley's finding that there was a bump in latency immediately preceding shortcut discovery was not found in this study. In this study, however, several problem type and size combinations were used, making the finding difficult to verify. Finally, Siegler and Crowley found that students never used incorrect solution procedures. In Study 2, some of the younger students overgeneralized the inversion-based shortcut and used it, incorrectly, on the standard problems. None of the older students ever misused the inversion-based shortcut on standard problems. Again, although the students in both grades did not spontaneously use the inversion-based shortcut, age differences appear to exist.

In Study 3, stability of the use of the inversion-based shortcut and the subtraction-first shortcut over time was assessed. Students who used the inversion-based shortcut frequently or infrequently in Study 1 were included in the analyses. Results supported the hypothesis that shortcut use would either remain stable for students who used the shortcut frequently or would generalize for students who used the shortcut infrequently on the pretest. Patterns of latency, accuracy, and self-report data were similar thereby providing more support for the veridicality of the self-reports. Again, age differences were found.

Although Users in both grades were equally likely to remain stable, Variable Users differed according to age. Variable Users in Grade 5 tended to become Users whereas Variable Users in Grade 1 were equally likely to increase, decrease, or maintain their shortcut use over time.

Performance on the evaluation-of-procedures task was at ceiling for the older students, whereas the younger students had difficulty recognizing and justifying the appropriate use of the shortcut. This finding also supports the hypothesis that students' reasons for shortcut use may differ according to age. Although the Grade 1 students in this study were capable of using the shortcut they did not seem to completely understand the principle of inversion which might indicate that they are using the shortcut out of necessity on the large problems.

The use of self-report data was very important for discovering that students use a subtraction-first shortcut based on the principle of associativity to solve standard problems. Self-report data also permitted closer examination on how students changed over time. Results from these studies are consistent with self-report data being veridical.

The results of these three studies indicate that although students in both grades performed similarly in Study 1, there are pervasive age differences. Performance on the evaluation-of-procedures task, precursors of discovery and generalization, inversion-based shortcut use according to problem size, and use of the subtraction-first shortcut all demonstrate these age-related differences. Age differences have important implications for theories of change as well as for instructional methods.

The age differences in precursors of shortcut discovery and generalization may have instructional implications. Exposure to an appropriate solution procedure appears to be an adequate instructional method for older students but direct instruction is necessary for some younger students. As well, the use of impasse problems should be considered as a useful tool for the discovery of new solution procedures for younger students. Development psychologists should be concerned with understanding the mechanisms that underlie developmental change (Kail & Bisanz, 1992). The results of Study 2 indicate that mechanisms of change are not the same for students of different ages, even though they may perform certain tasks equally. Therefore, models about change mechanisms should take these age-related differences into account.

References

Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. Developmental Review, 2, 213-236.

Ashcraft, M. H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In J. Bisanz, C. J. Brainerd, & R. Kail (Eds.), Formal methods in developmental psychology: Progress in cognitive development research (pp. 302-338). New York: Springer-Verlag.

Ashcraft, M. H. (1990). Strategic processing in children's mental arithmetic: A review and proposal. In D. F. Bjorklund (Ed.), Children's strategies: Contemporary views of cognitive development (pp. 185-211). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bisanz, J., & Lefevre (1990). Strategic and nonstrategic processing in the development of mathematical cognition. In D. F. Bjorklund (Ed.), Children's strategies: Contemporary views of cognitive development (pp. 213-244). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bisanz, J., Lefevre, J., & Gilliland, S. (1989, April). Developmental changes in the use of logical principles in mental arithmetic. Poster presented at the meeting of the Society for Research in Child Development, Kansas City, MO.

Bisanz, J., Lefevre, J., & Gilliland, S. (in preparation). Development of conceptual knowledge in arithmetic: The case of inversion.

Carpenter, T. P., & Moser, J. M. (1983). The acquisition of addition and subtraction concepts. In R. Lesh & M. Landau (Eds.), Acquisition of mathematics concepts and processes (pp. 7-44). New York: Academic Press.

Catan, L. (1986). The dynamic display of process: Historical development and contemporary uses of the microgenetic method. Human Development, 29, 252-263.

Cooney, J. B., & Ladd, S. F. (1992). The influence of verbal protocol methods on children's mental computation. Learning and Individual Differences, 4, 237-257.

Dhaliwal, G. (1989). Understanding of the inversion-based shortcuts in elementary school children. Unpublished B.Sc. Honours thesis, University of Alberta, Edmonton, Alberta, Canada.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. Psychological Review, 87, 215-251.

Kail, R., & Bisanz, J. (1992). The information-processing perspective on cognitive development in childhood and adolescence. In R. J. Sternberg & C. A. Berg (Eds.), Intellectual development (pp. 229-260). New York: Cambridge University Press.

Koshmider, J. W., & Ashcraft, M. H. (1991). The development of children's mental multiplication skills. Journal of Experimental Child Psychology, 51, 53-89.

Luria, A. R. (1978). The development of writing in the child. In M. Cole (Ed.), The selected writings of A.R. Luria (pp. 145-194). White Plains, NY: M. E. Sharpe.

Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. Memory and Cognition, 17, 759-769.

Siegler, R. S. (1976). Three aspects of cognitive development. Cognitive Psychology, 8, 481-520.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. Journal of Experimental Psychology: General, 116, 250-264.

Siegler, R. S. (1989). Hazards of mental chronometry: An example from children's subtraction. Journal of Educational Psychology, 81, 497-506.

Siegler, R. S., & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. American Psychologist, 46, 606-620.

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

Siegler, R. S., & Shipley, C. (1992). Variation, selection, and cognitive change. In G. Halford, & T. Simon (Eds.) Developing cognitive competence: New approaches to process modeling. Hillsdale, NJ: Lawrence Erlbaum Associates.

Siegler, R. S., & Shrager, J. (1984). A model of strategy choice. In C. Sophian (Ed.), Origins of cognitive skills (pp. 229-293). Hillsdale, NJ: Lawrence Erlbaum Associates.

Starkey, P., & Gelman, R. (1982). The development of addition and subtraction abilities prior to formal schooling in arithmetic. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), Addition and subtraction: A cognitive perspective (pp. 99-116). Hillsdale, NJ: Lawrence Erlbaum Associates.

VanLehn, K. (1988). Towards a theory of impasse-driven learning. In H. Mandl, & A. Lesgold (Eds), Learning issues for intelligent tutoring systems (pp. 19-41). New York: Springer-Verlag.

Vygotsky, L. S. (1978). Mind and society: The development of higher mental processes. Cambridge, MA: Harvard University Press.

Appendix A

Stimulus List (Order 1) for Pretest

2 + 7 - 2 =

8 + 6 - 4 =

3 + 24 - 26 =

9 + 27 - 27 =

8 + 4 - 8 =

6 + 27 - 26 =

4 + 8 - 6 =

26 + 3 - 26 =

5 + 29 - 27 =

2 + 4 - 5 =

8 + 6 - 6 =

7 + 21 - 21 =

3 + 9 - 7 =

2 + 25 - 25 =

6 + 2 - 3 =

24 + 5 - 24 =

8 + 26 - 28 =

5 + 7 - 7 =

4 + 23 - 25 =

22 + 6 - 22 =

9 + 7 - 5 =

4 + 28 - 28 =

3 + 4 - 4 =

Appendix B

Preliminary Briefing

"We are trying to find out what students know about different math problems. Can you help me with this, (student's name)? I will give you some problems to think about and then I will ask you some questions. Some of the problems you might find easy and some of the problems you might find hard. What I'm interested in is how you think about math so don't worry if you have trouble with some of the problems because this isn't a test. All I want you to do is to try your best. I will be writing down what you say so that I can remember what you said later on. I will also be keeping time with this stopwatch because some of the problems may take you longer than others and I want to know which ones they are. Do you have any questions before we start? Are you ready to start?"

Appendix C

Practice Problems

1 + 2 =

4 - 1 =

3 + 5 =

7 - 2 =

2 + 1 - 1 =

4 + 2 - 1 =

Appendix D

Discontinuation Protocol

If after 40 seconds, the child was unable to solve the problem or showed obvious signs of frustration, the experimenter asked "What are you trying to do?" If the child said he or she did not know, or appeared unable to solve the problem, the experimenter asked "If you could do this problem, what would you do first, and then what would you do?"

Appendix E

Stimulus Lists (Order 1) for Study 2

Session 1	Session 2	Session 3
4 + 5 - 5=	6 + 9 - 6 =	9 + 8 - 8 =
21 + 7 - 21 =	4 + 23 - 23 =	4 + 25 - 26 =
2 + 5 - 2 =	6 + 3 - 5 =	28 + 3 - 28 =
9 + 4 - 7 =	29 + 7 - 29 =	4 + 7 - 4 =
3 + 22 - 22 =	3 + 2 - 2 =	2 + 26 - 26 =
23 + 4 - 23 =	9 + 25 - 25 =	21 + 8 - 21 =
3 + 5 - 2 =	2 + 26 - 23 =	5 + 6 - 9 =
25 + 6 - 25 =	8 + 3 - 8 =	7 + 29 - 29 =
7 + 4 - 7 =	24 + 2 - 24 =	3 + 4 - 3 =
3 + 26 - 27 =	8 + 7 -7 =	6 + 2 - 2 =
8 + 3 - 3 =	6 + 28 - 28 =	24 + 9 - 24 =
5 + 27 - 27 =	4 + 8 - 7 =	2 + 7 - 6 =
8 + 6 - 8 =	5 + 2 - 5 =	9 + 23 - 22 =
9 + 24 - 24 =	4 + 27 - 28 =	5 + 9 - 9 =
6 + 28 - 25 =	26 + 5 - 26 =	8 + 24 - 24 =
2 + 6 - 6 =	7 + 6 - 6 =	6 + 8 - 6 =

Session 4	Session 5
3 + 6 - 3 =	6 + 8 - 8
8 + 29 - 29 =	23 + 5 - 23
4 + 5 - 6 =	6 + 28 - 25 =
28 + 4 - 28 =	4 + 2 - 4 =
5 + 24 - 22 =	7 + 22 - 22 =
2 + 9 - 9 =	8 + 3 - 6 =
6 + 23 - 23 =	27 + 9 - 27 =
8 + 5 - 8 =	9 + [illegible] - 4 =
25 + 3 - 25 =	3 + 28 - 28 =
5 + 26 - 26 =	7 + 8 - 7 =
4 + 3 - 3 =	2 + 7 - 5 =
9 + 23 - 25 =	25 + 8 - 25 =
5 + 9 - 5 =	5 + 2 - 2 =
8 + 4 - 2 =	4 + 23 - 26 =
27 + 2 - 27 =	2 + 24 - 24 =
7 + 5 - 5 =	9 + 3 - 9 =

Appendix F

Stimulus Lists for Evaluation-of-Procedures

Small Simple Inversion

2 + 7 - 7

3 + 9 - 9

8 + 5 - 5

7 + 4 - 4

5 + 2 - 2

6 + 3 - 3

4 + 8 - 8

9 + 6 - 6

Large Simple Inversion

5 + 22 - 22

8 + 27 - 27

4 + 28 - 28

7 + 25 - 25

3 + 29 - 29

6 + 23 - 23

2 + 26 - 26

9 + 24 - 24

Small Reversed Inversion

3 + 5 - 3

6 + 9 - 6

4 + 8 - 4

7 + 2 - 7

9 + 4 - 9

5 + 7 - 5

8 + 3 - 8

2 + 6 - 2

Large Reversed Inversion

21 + 3 - 21

26 + 7 - 26

22 + 9 - 22

28 + 4 - 28

27 + 2 - 27

25 + 6 - 25

23 + 8 - 23

24 + 5 - 24

Small and Large Standard

4 + 25 - 22

7 + 3 - 5

6 + 9 - 2

5 + 27 - 24

8 + 23 - 26

9 + 7 - 5

4 + 8 - 7

2 + 24 - 27

Appendix G

Instructions for Evaluation-of-Procedures Task

For each list of problems, each student was asked "What is the same about all of the problems in this list?" If the student did not know how the problems were alike, the following prompt was given: "How are all of these problems alike?" If the student still was not able to tell the experimenter how the problems were alike the experimenter said: "For all of these problems, you add one number and then take away (subtract) the same number." After the experimenter was satisfied that the student understood how the problems were alike the experiment asked the following question for both simple inversion lists: "A boy/girl I know says that if you start with a certain number (point to _a_) and you add another number (point at the first _b_) then take away that same number (point at the second _b_), the answer is always going to be the first number you started with (point to _a_). Would that way of solving this problem give you the right answer for all of these problems?" For both reversed inversion lists, the experimenter said: "A boy/girl I know says that if you start with a certain number (point to the first _b_) and you add another number (point to _a_) and then you take away the same number that you started with (point at the second _b_), then the answer is always going to be the middle number (point to _a_). Would that way of solving this problem give you the right answer for all of these problems? For the list of standard problems, the experimenter either gave the instructions for the simple inversion problems or for the reversed inversion problems.

After every question, each student was asked to justify their answer. The experimenter said "Why do you think that?" and if a prompt was needed, "Why would that work/not work?". If the student seemed to be trying to compute the answer for each problem, the experimenter asked: "What are you doing?" and if the student said that he or she was computing, the experimenter said "Do you need to do that?" and asked the student to justify his or her answer.

Appendix H

Extra Problems (Order 1) After Evaluation of Procedures

5 + 8 - 8 =

22 + 9 - 22 =

4 + 8 - 6 =

2 + 24 - 24 =

4 + 3 - 4 =

5 + 23 - 24 =

Appendix I

Direct Instruction 1

"I am going to show you a way to solve some of the problems that I have been giving you." Experimenter then showed list of two-term problems of the form b - b (Appendix L). The experimenter then went through each of the problems and then said "So, if you have a number and then you take away the same number what will always be the answer? Zero, that's right. Now, if I show you a problem like this (show 4 + 12 - 12), if you look at the problem carefully, you will see that there is a 12 and a take away 12 so what does 12 - 12 equal? (experimenter covered over the 4 on the sheet so the student only saw 12 - 12). Yes, that makes zero and what happens when you add 4 + 0? (experimenter covered the 12 - 12). Good, now what if you had a problem like 12 + 4 - 12? Well, there is still a 12 and a take away 12, isn't there? And what is 12 - 12 (experimenter covered the + 4). Okay, and what is 0 + 4? So, when you have a problem where you add and take away the same number (point to the b terms), the answer will always be the other number (point to the a). Now let's practice with four more problems (see Appendix L) and then I'll ask you to solve some problems on your own (see Appendix M)."

Appendix J

Direct Instruction 1 and 2 Session Problems

Instructional Problems

1 - 1 =

5 - 5 =

10 - 10 =

12 - 12 =

20 - 20 =

23 - 23 =

Practice Problems

2 + 23 - 23 =

4 + 6 - 6 =

8 + 3 - 3 =

27 + 5 - 27 =

Appendix K

Direct Instruction 1 Extra Problems

Order 1

8 + 7 - 8 =

5 + 28 - 28 =

7 + 22 - 24 =

3 + 6 - 6 =

24 + 3 - 24 =

6 + 7 - 5 =

Order 2

7 + 8 - 8 =

28 + 5 - 28 =

7 + 22 - 24 =

6 + 3 - 6 =

3 + 24 - 24 =

6 + 7 - 5 =

Appendix L

Direct Instruction 2

Direct instruction 2 was almost the same as direct instruction 1 except for two things: the experimenter used pennies to demonstrate all of the two-term problems and there was an extra list of two-term problems of the form $\underline{a}$ + 0 (see Appendix O). Practice problems were the same as in direct instruction 1 but the extra problems were different (see Appendix P).

Appendix M

Direct Instruction 2 Session Problems

1 + 0 =

4 + 0 =

8 + 0 =

11 + 0 =

17 + 0 =

25 + 0 =

Appendix N

Direct Instruction 2 Extra Problems

Order 1

3 + 5 - 2 =

29 + 4 - 29 =

5 + 23 - 23 =

7 + 4 - 4 =

8 + 23 - 26 =

6 + 2 - 6 =

Order 2

3 + 5 - 2 =

4 + 29 - 29 =

23 + 5 - 23 =

4 + 7 - 7 =

8 + 23 - 26 =

2 + 6 - 6 =

Appendix O

Stimulus List (Order 1) for Control Group

9 + 8 =

4 + 25 =

28 - 3 =

7 - 4 =

2 + 26 =

21 - 8 =

5 + 6 =

7 + 29 =

4 - 3 =

6 + 2 =

24 - 9 =

7 - 6 =

23 - 22 =

5 + 9 =

8 + 24 =

8 - 6 =

Table 1

Percentages of Discontinuations as a Function of Grade, Type, and Size in Study 1

	Grade 1			Grade 5		
	Small	Large	Mean	Small	Large	Mean
Simple Inversion (n=120 at each size)	29.2	43.4	36.3	0	0	0
Reversed Inversion (n=120 at each size)	25.8	45.0	35.4	0	0.8	0.4
Standard (n=240 at each size)	48.8	74.2	61.5	0	0.8	0.4
Mean (n=480 at each size)	38.1	59.2	48.6	0	0.6	0.3

Table 2

Percentages of Changed Answers as a Function of Grade, Type, and Size in Study 1

	Grade 1			Grade 5		
	Small	Large	Mean	Small	Large	Mean
Simple Inversion (n=120 at each size)	0	1.7	0.8	4.2	0	2.1
Reversed Inversion (n=120 at each size)	1.7	0	0.8	2.5	3.3	2.9
Standard (n=240 at each size)	0.4	0	0.2	2.1	2.9	2.5
Mean (n=480 at each size)	0.6	0.4	0.5	2.7	2.3	2.5

Table 3

<u>Mean Proportion Correct as a Function of Grade, Type, and Size in Study 1</u>

	Grade 1			Grade 5		
	Small	Large	Mean	Small	Large	Mean
Simple Inversion	.61	.52	.56	.93	.96	.95
Reversed Inversion	.63	.49	.56	.96	.93	.94
Standard	.32	.08	.20	.92	.82	.87
Mean	.52	.36	.44	.94	.90	.92

Table 4

Mean Latencies (in seconds) as a Function of Type and Size at Each Grade in Study 1

	Grade 1			Grade 5		
	Small	Large	Mean	Small	Large	Mean
Simple Inversion	23.66	24.65	24.16	4.75	5.14	4.95
Reversed Inversion	20.27	25.10	22.69	3.86	4.79	4.33
Standard	31.67	38.38	35.03	5.07	8.15	6.61
Mean	25.20	29.38	27.29	4.56	6.03	5.14

Table 5

Proportions of Self-Reported Solution Strategies as a Function of Type and Size for Grade 1 in Study 1

	Simple		Reversed		Standard		
	Small	Large	Small	Large	Small	Large	Total
L-to-R	.58	.40	.61	.44	.88	.79	.67
S	0	0	.01	0	.03	.04	.02
N	.05	.06	.06	.07	0	0	.03
I	.32	.49	.31	.43	.01	.02	.20
O	0	0	0	0	0	0	0
A	.05	.05	.02	.06	.08	.15	.08

L-to-R: left-to-right

S: subtraction-first shortcut

I: inversion-based shortcut

N: negation

O: other

A: ambiguous

Table 6

Proportions of Self-Reported Solution Strategies as a Function of Type and Size for Grade 5 in Study 1

	Simple		Reversed		Standard		
	Small	Large	Small	Large	Small	Large	Total
L-to-R	.57	.43	.66	.50	.85	.74	.67
S	0	0	.02	0	.14	.24	.10
N	.02	.02	0	0	0	0	.01
I	.42	.54	.32	.48	0	0	.22
O	0	.01	0	.01	.01	.01	.01
A	0	0	.01	0	.01	.01	.01

L-to-R: left-to-right

S: subtraction-first shortcut

I: inversion-based shortcut

N: negation

O: other

A: ambiguous

Table 7

Reliabilities of Solution Strategies as a Function of Type and Size

	Simple		Reversed		Standard		
	Small	Large	Small	Large	Small	Large	Total
L-to-R	.92	.88	.86	1.00	.98	.98	.96
S	-	-	1.00	1.00	1.00	.93	.97
N	.33	0	0	0	-	-	.14
I	.95	.93	.88	.96	-	0	.93
O	.33	-	1.00	0	.33	1.00	.44
A	-	-	.50	1.00	.33	0	.43

L-to-R: left-to-right

S: subtraction-first shortcut

I: inversion-based shortcut

N: negation

O: other

A: ambiguous

Table 8

Mean Proportion of Inversion-Based Shortcut Use as a Function of Grade, Sex, Type, and Size in Study 1

Grade 1

	Male			Female		
	Small	Large	Mean	Small	Large	Mean
Simple	.32	.58	.45	.32	.40	.36
Reversed	.35	.43	.39	.27	.43	.35
Mean	.33	.51	.42	.29	.42	.35

Grade 5

	Male			Female		
	Small	Large	Mean	Small	Large	Mean
Simple	.45	.47	.46	.38	.62	.50
Reversed	.35	.50	.43	.29	.47	.38
Mean	.40	.48	.44	.34	.54	.44

Table 9

Mean Proportion of Subtraction-First Shortcut Use as a Function of Grade and Size in Study 1

	Size		
Grade	Small	Large	Mean
1	.03	.05	.04
5	.14	.20	.17
Mean	.04	.17	.11

Table 10

Proportion Correct by User Groups as a Function of Grade, Group, and Type in Study 1

	Grade 1			
	User	Variable User	Non-user	Mean
Simple	.94	.60	.15	.56
Reversed	.88	.55	.24	.56
Standard	.21	.23	.16	.20
Mean	.68	.46	.19	.44

	Grade 5			
	User	Variable User	Non-user	Mean
Simple	.99	.95	.90	.94
Reversed	1.00	.96	.86	.94
Standard	.86	.91	.84	.87
Mean	.95	.94	.87	.92

Table 11

Mean Latencies (in seconds) as a Function of Group, Type, and Grade in Study 1

	Grade 1			
	User	Variable User	Non-user	Mean
Simple	8.92	21.96	39.41	23.43
Reversed	9.16	22.28	34.57	22.00
Standard	34.53	32.57	38.10	35.07
Mean	17.54	25.60	37.36	26.83

	Grade 5			
	User	Variable User	Non-user	Mean
Simple	3.19	4.18	7.60	4.99
Reversed	3.36	4.03	5.66	4.35
Standard	5.60	6.27	8.04	6.64
Mean	4.05	4.83	7.10	5.33

Table 12

Sessions for the Experimental and Control Groups in Study 2

	p-s	e-of-p	d.i. 1	d.i. 2	extra
Pretest	E,C				
Session 1	E				
Session 2	E				
Session 3	E,C*	E,C			E,C
Session 4	E		E		E
Session 5	E			E	E
Posttest	E,C	E,C			

p-s = problem-solving task

e-of-p = evaluation-of-procedures task

d.i. 1 = direct instruction 1

d.i. 2 = direct instruction 2

extra = extra problems at end of session

E = experimental group

C = control group

* - the control group received 16 two-term problems formed from the 16 three-term problems that the experimental group received

Table 13

Mean Proportion of Inversion-Based Shortcut Use as a Function of Test, Type, and Size for the Experimental and Control Groups in Grade 1, Study 2

	Grade 1 Experimental Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	0	0	0	.72	1.00	.28
Reversed	0	0	0	.61	1.00	.39

	Grade 1 Control Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	0	0	0	.20	.27	.07
Reversed	0	0	0	0	.07	.07

P.S.E. = problem-size effect (small - large)

Table 14

<u>Mean Proportion of Inversion-Based Shortcut Use as a Function of Test, Type, and Size for the Experimental and Control Groups in Grade 5, Study 2</u>

	Grade 5 Experimental Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	0	0	0	1.00	.93	-.07
Reversed	0	0	0	1.00	.93	-.07

	Grade 5 Control Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	0	0	0	.57	.76	.19
Reversed	0	0	0	.57	.62	.05

P.S.E. = problem-size effect (small - large)

Table 15

Mean Proportion Correct as a Function of Test, Type, and Size for Both Groups in Grade 1, Study 2

	Grade 1 Experimental Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	.28	0	-.28*	.83	1.00	.17*
Reversed	.39	0	-.39*	.89	.94	.05
Standard	.25	0	-.25*	.42	0	-.42*

	Grade 1 Control Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	.27	.13	-.14*	.73	.40	-.33*
Reversed	.33	.20	-.13*	.47	.13	-.34*
Standard	.27	.20	-.07	.50	.17	-.33*

Note. P.S.E. = problem-size effect, * $p < .0004$.

Table 16

Mean Proportion Correct as a Function of Test, Type, and Size for Both Groups in Grade 5, Study 2

	Grade 5 Experimental Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	.93	.87	-.06	1.00	.93	-.07
Reversed	.93	.80	-.13*	1.00	.93	-.07
Standard	.90	.80	-.10*	.83	.93	.10*

	Grade 5 Control Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	.86	.90	.04	1.00	1.00	0
Reversed	.90	.76	-.14*	.90	1.00	.10*
Standard	.93	.71	-.22*	.93	.83	-.10*

Note. P.S.E. = problem-size effect, * $p < .0004$.

Table 17

Mean Latencies (in seconds) as a Function of Test, Type, and Size for Both Groups in Grade 1, Study 2

	Grade 1 Experimental Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	36.3	40.0	3.7	4.9	6.8	1.9
Reversed	32.6	40.0	7.4*	10.1	8.0	-2.1
Standard	37.3	40.0	2.7	28.9	40.0	11.1*

	Grade 1 Control Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	38.6	40.0	1.4	19.0	27.2	8.2*
Reversed	29.4	38.2	9.2*	21.7	40.0	18.3*
Standard	33.8	38.0	4.2*	23.8	38.6	14.8*

Note. P.S.E. = problem-size effect, * $p < .0004$.

Table 18

Mean Latencies (in seconds) as a Function of Test, Type, and Size for Both Groups in Grade 5, Study 2

	Grade 5 Experimental Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	8.6	6.8	-1.8	1.6	1.5	-0.1
Reversed	5.2	6.7	1.5	1.2	1.9	0.7
Standard	6.7	10.2	3.5*	5.6	5.4	-0.2

	Grade 5 Control Group					
	pretest			posttest		
	Small	Large	P.S.E.	Small	Large	P.S.E.
Simple	4.5	11.2	6.7*	3.7	2.7	-1.0
Reversed	4.9	6.6	1.7	2.7	3.5	0.8
Standard	5.2	11.3	6.1*	5.0	10.2	5.2*

Note. P.S.E. = problem-size effect, * $p < .0004$.

Table 19

Sessions for Discovery, Frequent Use, and Consistent Use for Each Problem Type and Size Combination per Subject in Grade 1, Study 2

	Discovery				Frequent Use				Consistent Use			
Ss	SS	LS	SR	LR	SS	LR	SR	LR	SS	LS	SR	LR
102	2	1	4	1	2	1	5	1	5	1	5	1
105	4+	4+	4+	4+	PO	5	PO	5	?	5	?	5
121	4	4	4+	4	-	4	-	4	-	4	-	4
123	4+	4+	4+	4+	5	PO	-	PO	5	?	-	?
125	1	1	1	1	3	1	1	2	3	1	1	2
138	4+	4+	4+	4+	-	PO	PO	PO	-	?	?	?

Note. SS = small simple inversion problems, LS = large simple inversion problems, SR = small reversed inversion problems, and LR = large reversed inversion problems. Numbers in each column refer to the session during which the shortcut was discovered, used frequently, or used consistently. + refers to shortcut use during test that followed direct instruction in Sessions 4 or 5 (see Table 12). P = posttest and

- denotes cases in which Study 2 ended before a subject could meet criteria for frequent or consistent use.

Table 20

Sessions for Discovery, Frequent Use, and Consistent Use for Each Problem Type and Size Combination per Subject in Grade 5, Study 2

	Discovery				Frequent Use				Consistent Use			
Ss	SS	LS	SR	LR	SS	LS	SR	LR	SS	LS	SR	LR
506	3+	1	3+	1	5	1	4	4	5	4	4	4
513	1	1	1	1	2	1	1	1	2	1	1	1
530	3+	3+	3+	3+	4	4	4	4	4	4	4	4
543	3+	3+	3+	3+	4	4	4	4	4	4	4	4
545	3	2	2	2	3	3	3	3	5	3	3	3
549	3+	3+	3+	3+	4	4	4	4	4	4	4	4

Note. SS = small simple inversion problems, LS = large simple inversion problems, SR = small reversed inversion problems, and LR = large reversed inversion problems. Numbers in each column refer to the session during which the shortcut was discovered, used frequently, or used consistently. + refers to shortcut use during test that followed the evaluation-of-procedures task in Session 3 (see Table 12).

Table 21

Evaluation-of-Procedures Task on Session 3 and Posttest for the Grade 1 Experimental Group, Study 2

	Grade 1 Experimental Group			
	Session 3		Posttest	
Ss	Recognition	Justification	Recognition	Justification
102	FP	FP	FP	FP
105	R	O	FP	FP
121	FP	FP	R	FP
123	O	O	R	J
125	FP	FP	R	J
138	FP	O	FP	O

R - recognizer

FP - false positive

O - other

J - justifier

IJ - incomplete justifier

Table 22

Evaluation-of-Procedures Task on Session 3 and Posttest for the Grade 1 Control Group, Study 2

	Grade 1 Control Group			
	Session 3		Posttest	
Ss	Recognition	Justification	Recognition	Justification
101	O	O	O	J
107	R	IJ	R	J
112	FP	O	FP	O
120	FP	FP	FP	FP
124	FP	O	FP	O

R - recognizer

FP - false positive

O - other

J - justifier

IJ - incomplete justifier

Table 23

Evaluation-of-Procedures Task on Session 3 and Posttest for the Grade 5 Experimental Group, Study 2

	Grade 5 Experimental Group			
	Session 3		Posttest	
Ss	Recognition	Justification	Recognition	Justification
506	R	J	R	J
513	R	J	R	J
530	R	J	R	J
543	R	J	R	IJ
545	R	J	R	J
549	R	J	FP	FP

R - recognizer

FP - false positive

O - other

J - justifier

IJ - incomplete justifier

Table 24

Evaluation-of-Procedures Task on Session 3 and Posttest for the Grade 5 Control Group, Study 2

	Grade 5 Control Group			
	Session 3		Posttest	
Ss	Recognition	Justification	Recognition	Justification
514	R	O	R	J
521	R	IJ	R	J
531	R	J	R	J
539	R	J	R	O
541	FP	FP	R	J
542	R	J	FP	FP

R - recognizer

FP - false positive

O - other

J - justifier

IJ - incomplete justifier

Table 25

Mean Proportion of Inversion-Based Shortcut Use as a Function of Test and Size, Study 3

	pretest	posttest	mean
Small	.50	.72	.61
Large	.72	.82	.77
Mean	.61	.77	.69

Table 26

Mean Proportion of Subtraction-First Shortcut Use as a Function of Grade and Test, Study 3

	pretest	posttest	mean
Grade 1	.04	.10	.07
Grade 5	.28	.49	.39
Mean	.16	.30	.23

Table 27

Mean Proportion Correct as a Function of Grade, Test, and Size, Study 3

	pretest			posttest		
	Small	Large	Mean	Small	Large	Mean
Grade 1	.63	.49	.56	.65	.59	.62
Grade 5	.95	.94	.95	.95	.91	.93
Mean	.79	.72	.76	.80	.75	.78

Table 28

Mean Latencies (in seconds) as a Function of Test, Type, and Size for Grade 5, Study 3

	pretest			posttest		
	Small	Large	Mean	Small	Large	Mean
Simple	4.1	3.1	3.6	2.1	1.9	2.0
Reversed	3.3	4.0	3.7	2.5	2.4	2.5
Standard	4.8	7.0	5.9	4.3	4.4	4.4
Mean	4.1	4.7	4.4	3.0	2.9	3.0

Table 29

User Group Classification on Pretest and Posttest, Study 3

	Grade 1			
	Posttest			
Pretest	Users	Variable Users	Non-users	Total
Users	11	1	0	12
Variable Users	5	6	3	14
Total	16	7	3	26

	Grade 5			
	Posttest			
Pretest	Users	Variable Users	Non-users	Total
Users	11	2	1	14
Variable Users	9	4	0	13
Total	20	6	1	27

Table 30

Evaluation-of-Procedures Task for the Grade 1 Users, Variable Users, and Non-Users on the Recognition Component in Study 3

Classification	Users	Variable Users	Non-users	Total
Recognizers	7	3	1	11
False Positives	8	4	4	16
Other	1	0	0	1
Total	16	7	5	28

Table 31

Evaluation-of-Procedures Task for the Grade 1 Users, Variable Users, and Non-Users on the Justification Component in Study 3

Classification	Users	Variable Users	Non-users	Total
Justifiers	6	2	0	8
Incomplete Just.	5	2	1	8
False Positives	4	1	2	7
Other	1	2	2	5
Total	16	7	5	28