UNIVERSITY OF ALBERTA

ASSESSMENT OF STUDENTS' UNDERSTANDING OF THE MATHEMATICAL PRINCIPLE OF INVERSION

BY

DONALD J. MABBOTT ©

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

DEPARTMENT OF PSYCHOLOGY

Edmonton, Alberta

SPRING, 1993

ISBN 0-315-82224-4

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Donald Mabbott

TITLE OF THESIS: Assessment of Students' Understanding of the Mathematical Principle of Inversion

DEGREE: Master of Arts

YEAR THIS DEGREE GRANTED: Spring, 1993

D. Mabbott

6303-112 Street
Edmonton, Alberta
Canada, T6H-2E7

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 ASSESSMENT OF STUDENTS' UNDERSTANDING OF THE MATHEMATICAL PRINCIPLE OF INVERSION submitted by DONALD J. MABBOTT in partial fulfillment of the requirements for the degree of MASTER OF ARTS.

Jeffrey Bisanz

Gay Lee Bisanz

Thomas Kieren

Date: 10 December 1992

Abstract

The purpose of the present experiment was to examine (a) the conditions that influence students' performance on activities designed to assess understanding, (b) whether different forms of understanding are associated with abstract principles and component procedures, and (c) whether a developmental sequence exists in the acquisition of these different forms of understanding. To address these questions, I designed two tasks to assess three aspects of children's knowledge about the mathematical principle of inversion. Forty students from Grade 2 and 40 students from Grade 6 were assessed individually with the two tasks. A problem-solving task was used to examine students' use of specific procedures to solve a problem. A comprehension task consisted of an evaluation component and a justification component. With the evaluation component, students' ability to evaluate the correct use of procedures was examined. The justification component was used to assess students' ability to explain the rationale behind the use of specific procedures. To address the three original questions, three types of analyses were performed. First, analyses of performance on each task were conducted to determine the relative influence of different conditions. Examination of the results revealed that students performance on the three activities was influenced by different conditions. Specifically, schooling appeared to have a greater influence on the comprehension task compared to the problem-solving task, while previous experience with inversion appeared

to have a greater influence on the problem-solving task compared to the comprehension task. Second, relations between the tasks were examined to determine whether different forms of understanding are associated with abstract principles and component procedures. Specifically, patterns in performance across the three activities were identified, and relations among these activities were analyzed. Patterns associated with knowledge of component procedures were distinguished from patterns reflecting knowledge of abstract principles, thus providing evidence for different forms of understanding. Finally, Guttman Scaling was used to determine whether a developmental sequence exists in the acquisition of these different forms of understanding. Examination of the Guttman Scale indicated that the acquisition of abstract principles precedes the acquisition of component principles in the development of the understanding of inversion.

A Word of Thanks

to my supervisor, Dr. Jeffrey Bisanz, for his consistent guidance and counsel during the last two years.

to the members of my supervisory committee, Drs Gay Lee Bisanz and Thomas Kieren, for reviewing an earlier version of this manuscript, and for their helpful commencs and suggestions.

to Kathrine Robinson, for her helpful comments at all stages during the preparation of this thesis.

to Karen Hubbard and Jackie Leighton, for their help coding the data used to calculate reliability, and for their comments on an earlier version of this manuscript.

to Mike Thibodeau, Tanis Ryan, and Anita Cumbleton for their comments on an earlier version of this manuscript.

to my wife, Brenda Mabbott, whose constant love and care has supported me during the preparation of this thesis.

and most importantly, to God, who has provided me with the skills, the help of others, and the opportunities to learn that have enabled me to complete my program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Assessment of Students' Understanding of the Mathematical Principle of Inversion

A clear and comprehensive definition of understanding in mathematics is important for educators and researchers alike. Although there has been productive research focused on identifying processes that underlie mathematical computation (Ashcraft, 1982; Siegler & Shrager, 1984; Campbell & Graham, 1985; Campbell, 1987, 1989), little is known about the relation between computational processes and understanding (Bisanz & LeFevre, 1992; Bisanz, LeFevre, & Gilliland, 1992). For psychologists interested in providing a full account of remembering and problem solving, an integrated and detailed description of the relation between computational processes and the common notion of understanding is needed (Bisanz & LeFevre, 1992). There is also growing recognition among educators and researchers that more emphasis needs to be placed on helping children understand mathematics rather than just teaching them rote procedures (Cobb, Wood, Yackel, Nicholls, Wheatley, Trigatti, & Perlwitz, 1991; Diagnostic Mathematics Program, Alberta Education, 1990; Greeno, 1983; Schoenfeld, 1985). Comprehension in a domain includes knowledge of the logical or semantic principles that define the structures of the domain. When children are taught only to memorize mathematical facts, their knowledge is incomplete (Charles & Lester, 1982; Krulik & Rudnick, 1988).

Unfortunately, understanding is often poorly and inconsistently defined. For example, measures of children's understanding have not been considered in enough detail to be assessed behaviorally and therefore have played an insignificant role in instructional design (Greeno, 1983). In order for instructors to facilitate children's understanding of concepts and principles of mathematics, and for researchers to examine the relation between computational processes and understanding, it is important to establish specific theoretical characterizations of understanding.

Greeno (1983) argued that understanding must be thought of in terms of different forms, rather than as a single construct. A narrow definition or criterion would tend to make understanding an all-or-none phenomenon. Bisanz and LeFevre (1992) also suggested that different forms of understanding must be recognized. Specifically, they argued that the context in which understanding is assessed must be considered. For them, context refers broadly to task demands and materials that are used to evaluate understanding. Depending on the context, they concluded that an individual may show evidence for some forms of understanding but not others. Like Greeno, they suggested that no single context is definitive for assessing understanding.

Instead, they proposed a framework consisting of two dimensions, activity and generality, that jointly define a contextual space for assessing understanding. Assessment then consists of identifying profiles of understanding, across contexts, that reflect underlying individual differences or sequences in acquisition (Bisanz & LeFevre, 1992).

Although the activities used to assess understanding vary considerably in terms of the demands placed on an individual, Bisanz and LeFevre (1992) grouped them into three general classes: application of procedures to solve problems; explicit justification of procedures; and evaluation of the procedures. Application of procedures refers to the spontaneous use of a solution procedure that reflects, or is at least consistent with, a concept or principle appropriate for that problem. In tasks involving justification of procedures, a student must give an explanation of a concept or principle. Evaluation of procedures refers to a person's decision about the applicability and correctness of a particular solution to a problem. Bisanz and LeFevre suggested that each of the activities can be used independently to assess understanding.

If understanding can be differentiated into different forms, as suggested by Greeno (1983) and by Bisanz and LeFevre (1992), is there a developmental sequence in the acquisition of these different forms of understanding? Currently there is little consensus about the nature of development of understanding in mathematics. The development of understanding has been described as either (a) a process in which abstract principles guide the acquisition and use of procedures (Bullock, Gelman, & Baillargeon, 1982; Gelman & Meck, 1983; Flavell, 1979; Silver, 1985; Starkey & Gelman, 1982) or (b) a process in which the acquisition and use of component procedures leads to the acquisition of abstract principles (Anderson, Boyle, & Yost, 1985; Baroody & Ginsburg, 1986; Fuson & Hall, 1983; Mayer, 1985).

The characterization of principles guiding the acquisition of procedures has been used to describe a number of different mathematical skills in children. Starkey and Gelman (1982) argued that a set of arithmetic principles underlies preschoolers' performance of addition, subtraction, inversion, and compensation tasks. Gelman (Gelman & Gallistel, 1978; Gelman & Meck, 1983) also suggested that knowledge of counting principles forms the basis for the acquisition of counting skills.

The idea that the acquisition and use of component procedures leads to the acquisition of abstract principles is based on the assumption that the repeated use of procedures, and repeated presentation of materials, results in conceptual associations that lead to the development of understanding (Anderson, Boyle, & Yost, 1985; Mayer, 1985). Further, increased understanding is seen as leading to procedural advances and more sophisticated application of procedures, resulting in more advanced principles (Baroody & Ginsburg, 1986). For example many researchers have argued that children apparently first learn to use numbers mechanically and only gradually construct an understanding of number and counting (Baroody & Ginsburg, 1984; Fuson & Hall, 1983).

RATIONALE

The present experiment was designed to examine (a) the conditions that influence students' performance on activities designed to assess understanding, (b) whether different forms of understanding are associated with abstract principles and component procedures, and (c) whether a developmental sequence exists in the acquisition of these different forms of understanding. To address these questions, I have designed two tasks to assess three aspects of children's knowledge about the mathematical principle of inversion. Bisanz and LeFevre (1990) argued that the effective use of shortcuts may serve as an operational measure of how well an individual functionally understands the concepts underlying arithmetic and the number systems. One such shortcut is based on the logical principle of inversion. In arithmetic problems that have the form $a + b - b$, the successive retrieval of a sum and difference can be very inefficient compared to the shortcut based on inversion. Using the principle of inversion, one can state the answer for inversion problems is a without doing any computation (because b - b is zero). The inversion principle has been used to study children's specific use of different mathematical principles (Bisanz, LeFevre, & Gilliland, in preparation; Starkey, & Gelman, 1982) but never to assess developmental sequences involved in the acquisition of understanding.

The two tasks are based on the activities proposed by Bisanz and LeFevre (1992). A problem-solving task was designed to examine the procedures students use spontaneously to solve a set of 3-term addition and subtraction problems (application-of-procedures activity). Two types of problems were presented, with large and small instances of each. Inversion problems (e.g., 5 + 7 - 7 = ?) could be solved using an inversion-based shortcut. Standard problems (e.g., 3 + 6 - 4 = ?) could be solved only by using some combination of addition and subtraction.

To examine the conditions that influence performance on the problem-solving task, students' accuracy, responses latencies, and self-reports about how they solved the problem were recorded. In mental arithmetic, accuracy tends to decrease, and response latencies tend to increase, on problems with larger numbers (the "problem-size effect") (Koshmider, & Ashcraft, 1992). This problem-size effect reflects the greater difficulty of computation with large numbers commonly found in research with children and adults (Koshmider, & Ashcraft, 1992). Consequently, accuracy and response latencies should always show a problem-size effect on Standard problems because these problems can be solved only by performing addition and subtraction. Students who do not use a shortcut also should show an identical problem-size effect on Inversion problems. In contrast, students who use a shortcut on Inversion problems should show no such effect because they do not use addition and subtraction to find solutions.

Students' self-reports were used to supplement the accuracy and latency data by providing additional insights about the procedures students used. Specifically, students were grouped according to the proportion of times they used inversion, and comparisons of accuracy and latency were made between these groups. Students who reported use of an inversion-based shortcut on inversion problems (i.e. students who said that the last two numbers equalled zero (b - b = 0), so the answer had to be the first number (a + 0)) would be expected to have high accuracy rates, and fast latencies on inversion problems. Students' accuracy, response latencies and self-reports were also used to examine changes in performance as children progress through school. Because of their greater experience with mathematics, older students would be expected to have higher rates of accuracy, have faster latencies, and report using inversion more often than younger students.

The <u>comprehension task</u> had two components: the <u>evaluation component</u> corresponds to the evaluation-of-procedures activity described by Bisanz and Lefevre (1992), and the <u>justification component</u> corresponds to the justification-of-procedures activity. Students were shown 3 sets of 3-term addition and subtraction problems. One set consisted of small inversion problems, 1 set consisted of large inversion problems, and 1 set consisted of standard problems. In the evaluation component students were asked to identify the appropriate use of a shortcut based on inversion to solve the problems for each of the 3 sets. This component was designed to address whether students can recognize the appropriateness of procedures based on abstract principles without requiring that the students be able to produce those procedures. In the justification component students were asked to justify their answer to the evaluation component. This component was designed to assess whether students can explain abstract principles even if they do not use procedures. For both of the components, students' self-reports were recorded. If, on inversion problems the student identified the appropriate use of a shortcut and explained why that shortcut worked, they were considered to have a conceptual understanding of the principle of inversion. Because of their greater experience with mathematics, older students would be expected to identify the appropriate use of an inversion-based shortcut and explain the principle of inversion verbally more often than younger students.

To evaluate the hypothesis that different forms of understanding co-exist, students' response patterns across the different activities were examined. Bisanz and LeFevre (1992) identified many problems associated with assessment based solely on one of the activities. For example, a student's spontaneous use of a shortcut on the problem-solving task may appear to be compelling evidence that he or she understands the principle of inversion. However, a student may use a conceptually appropriate procedure for reasons unrelated to the underlying concept. Similarly, failure to provide an adequate justification does not imply that a student lacks knowledge about a concept. He or she may simply have difficulty verbalizing that knowledge. Conversely, being able to provide adequate explanations or rationales does not necessarily imply that the student can use the corresponding procedures spontaneously.

Because of the problems discussed above, I will examine relations among the activities and make inferences about possible forms of understanding associated with them. Specifically, students were classified as successful or unsuccessful on each activity (success was determined by the presence of accurate knowledge about inversion in students' self-reports). Response patterns across the problem-solving task and the two components of the comprehension task were examined. If understanding of inversion is not differentiated into distinct forms, students would either successfully perform, or unsuccessfully perform, all the activities. However, if response patterns across the activities demonstrate that students respond successfully on some but not all activities, we may conclude that different forms of understanding are present.

Patterns of responses across these activities for individual students should be informative about the developmental sequence(s) in which understanding is acquired. For example, if principles are acquired first, followed by procedures, then students should perform accurately on the justification component, but not on the problem-solving task. If procedures are acquired first, followed by principles, students should perform accurately on the problem-solving task but not on the justification component.

METHOD

Subjects

Forty students from Grade 2 and 40 students from Grade 6 were assessed individually with the problem-solving task and the two components of the comprehension task. Previous researchers have indicated that few differences in children's overt performance exist between Grades 1 and 4 (Bisanz & LeFevre, 1989). Between Grades 2 and 6 a dramatic increase in the spontaneous use of inversion has been reported (Bisanz & LeFevre, 1989). Consequently, studying students in Grades 2 and 6 allows for the examination of a period where there may be an increase in the use of procedures based on inversion. Because only a small percentage (approximately 20%) of Grade 2 students have been observed to use inversion spontaneously (Bisanz & LeFevre, 1989; Dhaliwal, 1989) it is important to have a sample size large enough to observe any effects.

Materials

Problem-Solving Task

For the problem-solving task 16 problems were presented, including 8 inversion problems ($a + b - b = a$) and 8 standard problems ($a + b - c = d$). Inversion problems of the form $a + b - b$ (as opposed to $b + a - b$) were used because in previous studies students identified problems more easily when the identical numbers were in close proximity (Dhaliwal, 1989). Each problem type was divided into large and small instances to examine whether inversion-based strategies are employed more frequently on problems consisting of large numbers or problems consisting of small numbers (Dhaliwal, 1989). In the large form, "a" was a number between 1 and 9, and "b" and "c" were numbers between 21 and 31. In the small form, "a" was a number between 1 and 9, and the values of "b" and "c" were between 1 and 9. Half the intermediate sums (a + b) for each combination of problem type and problem size crossed decade boundaries, and half did not.

The order of the problems was generated unsystematically with a few constraints: All the combinations of size and type were spaced evenly throughout the order, and no problems with similar answers were placed beside each other. A second order was created by reversing the first. The complete set of problems is listed in Appendix A.

Comprehension Task

For the comprehension task 3 different sets of 3-digit addition and subtraction problems were presented. One set included 8 small inversion problems, one included 8 large inversion problems, and one included 8 standard problems. All of the problems presented were similar to those in the problem-solving task. Each set of problems was presented to each student on cue cards. The three sets of problems are listed in Appendix B.

Procedure

Students were assessed individually in two sessions, each lasting 15 minutes. Performance was recorded with a video camera to enable careful analysis of responses. Specifically, when the data were analyzed, the experimenter referred to the video tapes, as well observations recorded with pencil and paper. Each session began with a general briefing followed by specific instructions for each task. The full text for the briefing and instructions is provided in Appendix C. For the problem-solving task, the order of problem presentation was reversed to create a second order, and the two orders were counterbalanced across subjects. To evaluate whether previous exposure to content related to the principle of inversion influenced students' performance, task order was varied across subjects, with half receiving the problem-solving task first and half receiving the comprehension task first.

Problem-Solving Task

For the problem-solving task, students were asked to solve each problem and then state the answer. Students' response latencies were recorded with a stopwatch. They were then asked how they solved each problem. If after approximately 40 seconds a student was unable to solve the problem, or if they showed clear signs of frustration, the experimenter would prompt the student. For example the experimenter would ask the student what numbers they were looking at, or ask them how they were trying to solve the problem. Students were not asked to think aloud as they solved each problem because concurrent self-reports would have resulted in inflated responses latencies.

Comprehension Task

For the comprehension task, the experimenter described an inversion-based strategy that could be used to solve the problems in each set. He then asked the student if the strategy just described would work for the whole set of problems (evaluation component). After the student had given an answer, he or she was asked why the strategy would work or why it would not work (justification component). The text for each description is provided in Appendix C.

RESULTS AND DISCUSSION

Data from the problem-solving and comprehension tasks were analyzed in two ways. First, measures of performance within each task were analyzed to determine the effects of grade, sex, task order, problem size (small and large), and problem type (standard and inversion). Second, the relations between the tasks were examined to determine whether performance differed on measures of problem-solving, justification, and evaluation, and whether these three aspects of understanding form a developmental sequence.

Within-Task Analyses

Problem-Solving Task

Accuracy

Initially, the effects of schooling, sex, task order, problem size, and problem type on the proportion of accurate responses were examined. Determining what to score as an accurate response presented a few problems. Some students gave an answer and then, after discussing how they got that answer, they changed it. Other students found it difficult to solve some of the problems. If, after approximately 40 seconds, a student was unable to solve the problem, or if the student showed clear signs of frustration, the experimenter would use prompts to help him or her. To deal with these problems, accuracy was scored with two sets of criteria. Initial accuracy was based on a student's very first response. Answers of students who were prompted were coded as inaccurate because they were unable to give an answer spontaneously. Final accuracy included any changes in students' answers after discussion of the problem. Proportions of accurate responses for each method of scoring were then subjected to separate 2(Grade) X 2(Sex) X 2(Task Order) X 2 (Problem Size) X 2 (Problem Type) analyses of variance with repeated measures on the last two variables. Results of the two analyses were similar, the only difference being that a sex effect was found when analyzing final accuracy. Because of the similarities between the two analyses, and because the students' initial answers give a better indication of their spontaneous use of strategies, only the results for initial accuracy are presented in detail.

Students were more accurate on small than large problems (84.0% vs. 68.0%), $F(1, 72) = 51.36$, $p < .01$. This problem-size effect reflects the greater difficulty of computation with large numbers commonly found in research with children and adults. Accuracy was higher for inversion than for standard problems (88.0% vs. 65.0%), $F(1, 72) = 97.16$, $p < .01$, a result that would be expected if students were using the inversion-based shortcut on inversion problems. Finally, Grade 6 students were more accurate than Grade 2 students (90.0% vs. 62.5%), $F(1, 72) = 49.15$, $p < .01$.

Problem size interacted with grade, $F(1, 72) = 9.23$, $p < .01$. Mean rates of accuracy are listed in Table 1. An examination of the simple effects confirmed that Grade 6 students were more accurate that Grade 2 students for both problem sizes, $Fs(1, 72) > 22.20$, but this difference was greater for large problems. Although students in both grades were more accurate on small versus large problems, $Fs(1, 72) > 8.42$, $ps < .01$, this difference was greater for Grade 2 students. Thus, the greater computational demands of large problems appears to have been especially difficult for students in Grade 2. These data should be interpreted cautiously however, because of the high accuracy rates for Grade 6 students. With high accuracy rates there is substantial possibility of ceiling effects, which might render the statistical comparisons misleading.

Problem size also interacted with problem type, $F(1, 72) = 28.99$ $p < .01$. Mean rates of accuracy are provided in Table 2. An examination of the simple effects confirmed that students were more accurate on inversion than standard problems for both problem sizes, $Fs(1, 72) > 26.08$, $ps < .01$, but this difference was greater for the large problems. These results are consistent with the conclusion that students in both grades used a shortcut on Inversion problems, at least to some extent. Students were more accurate on small versus large problems for both standard and inversion problems $Fs(1, 144) > 6.46$, $ps < .05$. However, the difference between large and small was greater for standard problems than for inversion problems. Because of the greater computational demands of large problems, when students employ a successive addition and subtraction strategy, accuracy decreases as problem size increases (problem-size effect). As expected, this effect was especially evident for standard problems, where an inversion-based shortcut could not be used. If students were using successive addition and subtraction on inversion problems, we would expect the difference between large and small problems to be the same for both standard and inversion types. Because the difference was smaller for inversion problems, the use of a shortcut may moderate the problem-size effect here. Again, these data should be interpreted cautiously because of the substantial possibility of ceiling effects.

Interactions between problem type and grade, $F(1, 72) = 23.26$, $p < .01$, and problem type and task order, $F(1, 72) = 13.14$, $p < .01$, were qualified by the interaction between type, grade, and task order, $F(1, 72) = 8.88$, $p < .01$. Means are provided in Table 3. An examination of simple effects revealed that students in both grades were more accurate on inversion problems than standard problems in both task orders, $Fs(1, 72) >$

4.82, *ps* < .05. This difference was largest for Grade 2 students who had previously been asked to recognize and justify inversion (Comprehension First). The data are consistent with the conclusion that previous exposure to the inversion procedures increased the probability that students in Grade 2 used a inversion-based shortcut on the problem-solving task. Use of a shortcut may in turn yield higher rates of accuracy because the calculational demands of these problems are minimized when an inversion procedure, as opposed to successive addition and subtraction, is used. In Grade 2, students who started with the problem-solving task were more accurate on standard problems than those who started with the comprehension task, $F(1, 144) = 5.29$, $p < .05$. Students who had the comprehension task first were more accurate on inversion problems than those who had the problem-solving task first, $F(1, 144) = 55.39$, $p < .01$. Even though the Grade 2 students who started with the comprehension task were not as proficient in solving standard problems as those who started with the problem-solving task, their previous experience with the principle of inversion on the comprehension task was sufficient to increase their performance on inversion problems to the point where their accuracy surpassed that of the other students.

No differences in accuracy were found between task orders for Grade 6 students, $Fs(1, 144) < 1$. Presumably, the problems were so easy for the Grade 6 students that previous experience was unnecessary for increased accuracy. Grade 6 students in both task orders were more accurate than Grade 2 students in both task orders on standard problems, $Fs(1, 144) > 21.69$, $ps < .01$. As well, on inversion problems, Grade 6 students who had the problem-solving task first were more accurate than Grade 2 students who started with the problem-solving task, $F(1, 144) = 11.45$, $p < .01$. There was no difference between students in the two grades on inversion problems for students who had the comprehension task first, $F(1, 144) = 2.50$, $p < .10$. Thus the effect of previous experience is large enough to minimize grade differences in performance on inversion problems.

Based on the analyses of accuracy data, it appears that students used a shortcut based on inversion, and that previous experience with that shortcut may improve accuracy considerably. Again, these data should be interpreted cautiously because the high accuracy rates suggest ceiling effects, which might render the statistical comparisons misleading. Subjects' response latencies are likely to be a more sensitive measure of existing effects.

Latency

Effects of schooling, sex, task order, problem size, and problem type on latency were examined. Using median latencies, as opposed to mean latencies, has the advantage of eliminating extreme responses that might be due to extraneous factors.

Median latencies for responses were computed for each combination of problem type and problem size, generating 4 medians (2 problem sizes X 2 problem types) per subject. Response latencies for problems where the student was inaccurate might reflect different processing than for problems with accurate performance. Therefore, only the latencies on problems with accurate performance were used in the analyses of latency. As mentioned earlier, students sometimes gave an answer, and then after some discussion of how they got the answer, they would change it. In these instances, I was unable to compute latencies for the problems. This situation occurred infrequently however (4.80% of problems for Grade 2, and 3.75% of problems for Grade 6). On problems where students were prompted by the experimenter, latencies were estimated to be 40 seconds. Because the variability in latencies was much greater for Grade 2 students than Grade 6 students separate ANOVAs were computed at each grade level. For each grade, median latencies were subjected to a 2(Sex) X 2(Task Order) X 2(Problem Size) X 2(Problem Type) analysis of variance, with repeated measures on the last two variables.

For Grade 2, latency varied as a function of problem size and problem type. Students had faster latencies for small problems than for large problems (9.1 vs. 18.5 s), $F(1, 30) = 22.57$, $p < .01$, and faster latencies for inversion problems than standard problems (8.9 vs. 18.7 s), $F(1, 30)$ 45.97, $p < .01$. As well, problem size interacted with problem type, $F(1, 30) = 9.49$, $p < .01$ (see Figure 1). An examination of simple effects revealed that students had faster latencies for inversion problems than standard problems for both sizes, $Fs(1, 60) > 6.50$, $ps < .05$. For both standard and inversion types, students had faster latencies for small problems than large problems $Fs(1, 60) > 11.28$, $ps < .01$. If students were using successive addition and subtraction to solve both standard and inversion problems, the difference between small than large problems should be the same for both types. However the difference between small and large was greater for standard problems (12.2 vs. 25.2 s) than for inversion problems (6.0 vs. 11.7 s). If all students were using a shortcut on inversion problems, we would expect there to be no difference between small and large problems. These results are consistent with the conclusion that Grade 2 students used shortcuts on inversion problems some of the time.

Grade 6 students also had faster latencies for small problems versus large problems (3.4 vs. 5.9 s), $F(1, 36) = 35.46$, $p < .01$, and faster latencies for inversion problems than standard problems (3.0 vs. 6.3 s), $F(1, 36) = 45.63$, $p <.01$. Again, problem size interacted with problem type, $F(1, 36) = 25.10$, $p < .01$ (see Figure 2). An examination of simple effects revealed that for small problems there was no clear difference in the latencies for standard or inversion problems, $F(1, 72) = 3.44$, $p > .05$. Presumably, the small problems were so easy that, even if students had used shortcuts on inversion problems, solution latencies were no faster than when successive addition and subtraction was used on standard problems. For large problems however, students had faster latencies on inversion problems than standard problems, $F(1, 72) = 70.36$, $p < .01$, indicating they used a shortcut on large inversion problems. Consistent with the results for Grade 2 students, Grade 6 students had faster latencies on small than on large standard problems, $F(1, 72) = 60.02$, $p < .01$. If students were using successive addition and subtraction on all problems, an increase in latency between small and large inversion problems would be expected. However, there was no difference in students latencies between small and large inversion problems, $F(1, 72) < 1$. These results are consistent with the conclusion that students were using a shortcut on inversion problems.

For Grade 6, problem type also interacted with task order, $F(1, 36) = 6.87$, $p < .05$, and sex, $F(1, 36 = 6.14$, $p < .05$. Examination of the simple effects for the type by task order interaction indicated that students in both task orders had faster latencies on inversion problems than standard problems, $Fs(1, 36) > 8.54$, $ps < .01$. However, the difference between inversion and standard problems was emphasized for students who did the comprehension task first (2.4 vs. 7.0s for comprehension first, and 3.7 vs. 5.7s for problem-solving first). This difference may be due to the fact that students who started with the comprehension task had previous experience with inversion principles. Examination of simple effects for the type by sex interaction revealed that both boys and girls were faster on inversion problems than standard problems, $Fs(1, 36) > 9.14$, $ps < .01$, but the difference between the two problem types was greater for girls (2.9 vs. 7.4s) than for boys (3.2 vs. 5.3s).

Taken together, the accuracy and the latency analyses indicate that students used shortcuts on inversion problems, at least to some extent.

Self-Reports

To evaluate the nature of the strategies students used, their self-reports were examined. Examination of self-reports provided converging evidence with the latency and accuracy data about students use of shortcuts based on inversion. Comparing students' self-reports with the accuracy and latency data also allowed for evaluation of the verdicality of those self-reports. Specifically, reports were classified into different categories based on strategy use, the reliability of these classifications was assessed, and the frequency of use for each of the strategies was calculated. The reports were then used to interpret the accuracy and latency data. Self-reports served as a verbal confirmation of shortcut use.

Students' responses were classified using two levels of analysis. First, each response was grouped according to the strategy used by the student to solve the problem (e.g., left-to-right addition and subtraction, or inversion-based strategy). Then each response was classified according to the specific steps that students may have used to carry out their strategy (e.g. for left-to-right addition and subtraction they may have counted on their fingers, or used a derived fact). In classification, the fact that students may have used more than one step, or no visible steps, was taken into account. A description of each of the strategies observed and the procedures for scoring are found in Appendix D.

An independent rater classified responses from 20% of the subjects, and reliability with the experimenter's classification was calculated for each combination of problem type and problem size. Specifically, the number of agreements between the rater and the experimenter was determined, and this number was divided by the total number of responses. Reliabilities for the small inversion, large inversion, small standard, and large standard problems were .91, .94, .78, and .81, respectively. Reliabilites were low for the standard problems because the diverse and ambiguous responses students gave on these problems were hard to classify. Any conclusions based on analyses of the standard problems must be qualified by the low raliabilities.

Examination of how frequently each strategy was used revealed three major categories: inversion-based strategies, left-to-right addition and subtraction strategies, and subtraction-first strategies. Inversion-based strategies were used only on inversion problems, and subtraction-first strategies were used only on standard problems. Left-to-right strategies were used most frequently, and were used on both problem types. Tables 4 and 5 detail the relative frequency of each strategy for every combination of problem type and problem size.

Inversion problems. The proportion of times students reported using inversion-based shortcuts was calculated for large and small problems. Proportion of inversion use was then subjected to a 2(Grade) X 2 (Sex) X 2(Task order) X 2(Problem Size) analysis of variance with repeated measures on the last variable. Somewhat surprisingly, given the amount of arithmetic instruction in school, frequency of use increased only slightly from ages 7 to 11 (38% vs. 55%). Because of substantial variability within each age group this difference was not significant, $F(1, 72) = 3.52$, $p = .07$. Students reported the use of inversion shortcuts more frequently on large problems than small problems (49.7% vs. 43.8%), $F(1, 72) = 6.10$, $p < .05$. This result is consistent with the conclusion that on larger problems, where students have a harder time getting the answer by simple addition and subtraction, they are more likely to use an alternative, easier strategy. Finally, students used the shortcut more when they had previous exposure with inversion on the comprehension task (comprehension first), than when they did not (problem-solving first) (60.0% vs. 33.5%), $F(1, 72) = 8.40$, $p < .01$. These results are consistent with the conclusion that students are more likely to use inversion spontaneously if they have had previous exposure to it.

Because of the substantial variability within each age group, the means for each group may not have accurately represented the distribution of scores. Consequently, the number of students who used inversion 0, 1, 2, 3, and 4 times was calculated for each grade and each problem size and these distributions were then analyzed using nonparametric statistics. When analyzed with the Kolmogorov-Smirnov two-sample test, no differences between the distributions of inversion use were evident between Grade 2 and 6 for small problems, K-S $Z = .894$, $p > .30$, and for large problems, K-S $Z = .783$, $p > .50$. When the same distributions were analyzed with the chi-square two-sample test, no differences between the grades were evident for small problems, $X^2(4, N = 80) = 3.06$, $p > .50$, and for large problems, X^2 $(4, N = 80) = 4.02$, $p > .30$. These results are consistent with the analysis of variance. Although some students in both grades used inversion spontaneously, there was little increase in its use from Grades 2 to 6.

To examine the verdicality of self-reports, as well as to identify converging evidence, accuracy and latency data were re-evaluated using self-reports. Students were categorized into 3 groups based on the frequency of times they reported using inversion. Students who reported using inversion frequently would be expected to show different accuracy and latency patterns than other students. Students were categorized as (a) inversion non-users if they did not use inversion on either the small or large inversion problems, (b) infrequent inversion users if they used inversion between 1 and 5 times on both small and large problems, and (c) frequent inversion users if they used inversion between 6 and 8 times on both large and small problems. In Grade 2 there were 18 inversion non-users, 8 infrequent users, and 14 frequent users. In Grade 6 there were 11 non-users, 9 infrequent users, and 20 frequent users. Both accuracy and latency data were reanalyzed, and inversion use was included as a between-subjects variable. The proportion of accurate responses was subjected to a 2(Grade) X 2(Sex) X 2(Task order) X 3(Inversion use) X 2(Problem Type) X 2(Problem Size) analysis of variance with repeated measures on the last two variables. Only the effects that involved the inversion use categories are discussed.

An interaction between inversion use and problem type, $F(2, 56) = 3.94$, $p < .05$, was qualified by the interaction between grade, inversion use and problem type, $F(2, 56) = 4.24$, $p < .05$ (see Table 6). Although all Grade 2 students were more accurate on inversion versus standard problems, $Fs(1, 112) > 10.4$, $ps < .01$, students who reported using a shortcut frequently showed the greatest difference in accuracy between the two problem types (.99 vs. .41). The greater accuracy of frequent users can be attributed to their use of an inversion-based shortcut. However, the accuracies of students who did not report using a shortcut, or reported using one infrequently, were also higher for inversion problems. These students may have used a shortcut that they did not mention. For example, infrequent users and non-users may have used the strategy of negation (Bisanz & LeFevre, 1990), which may have increased their accuracy. Inversion problems may just have been easier than standard problems. A similar pattern was observed for Grade 6 students, although the difference between standard and inversion problems was only significant for students who reported using a shortcut frequently, $F(1, 112) = 6.66$, $p < .05$. Students who said they used an inversion-based strategy to solve inversion problems had higher rates of accuracy on inversion than standard problems. As noted earlier, these results should be interpreted cautiously because of possible ceiling effects.

Grade 2 students who frequently used a shortcut were just as accurate as Grade 6 students on inversion problems, $F(1, 112) <$

1. Because of this result, we might expect that greater proficiency in calculation, compared to their fellow students, accounted for frequent use of the shortcut. If this were the case we would expect frequent users to perform more accurately on all kinds of problems. Although Grade 2 frequent users were more accurate than non-users, and infrequent users on inversion problems, Fs(1, 112) > 11.5, ps < .01, they were not more accurate than the others on standard problems Fs(1, 112) < 1.94, ps > .10. Because Grade 2 frequent users were not more proficient than other students in all their calculations, other variables must be considered to account for their use of the shortcut. Perhaps frequent users were more flexible in their thinking when solving problems. This hypothesis is considered in more detail in a later section. In Grade 6 there was no difference in accuracy between the three categories for both standard and inversion problems, Fs < 1.

Inversion use also interacted with problem size, $F(2, 56) = 4.15$, $p < .05$ (see Table 7). Because of the computational demands of large problems, infrequent users and non-users should be more accurate on small problems compared to large problems (i.e. the problem-size effect). No differences between the two problem sizes would be expected for students who frequently used an inversion-based shortcut. Furthermore, students who reported frequent use of shortcuts should be more accurate on large problems compared to students who reported infrequent or no shortcut use. As expected, all three groups of students were equally accurate on small problems, *F*s(1, 112) < 0.30, *p*s > .25, and students who reported using shortcuts frequently were more accurate on large problems than those who reported little or no shortcut use, *F*s(1, 112) > 5.33, *p*s < .05. As well, infrequent users and non-users were more accurate on small compared to large problems, *F*s(1, 56) > 25.8, *p*s < .01, and frequent users were equally accurate on both problem sizes, $F(1, 56) = 3.26$, $p > .05$. These results are consistent with the conclusion that some students use a shortcut on inversion problems and that shortcut is based on the principle of inversion. As mentioned earlier, these results need to be interpreted with caution because of possible ceiling effects. A better indication is given by analyses of latency data.

For each grade, median latencies were subjected to a 2(Grade) X 2(Sex) X 2(Task Order) X 3(Inversion use) X 2(Problem Size) X 2 (Problem Type) analysis of variance with repeated measures on the last two variables. For Grade 2, inversion use interacted with problem type, $F(2, 22) = 6.62$, $p < .01$ (see Table 8), and with problem size, $F(2, 22) = 5.35$, $p < .05$ (see Table 9). Because the problems were relatively easy for Grade 6 students, all their latencies were fast, and no differences between the groups were significant.

If students who frequently reported using an inversion-based shortcut actually employed one, their latencies should be faster for inversion problems compared to standard problems. Students who reported infrequent or no use of a shortcut should have the same latencies for both problem types. Grade 2 frequent users and non-users had faster latencies on inversion than standard problems, $\underline{F}s(1, 22) > 14.10$, $\underline{p}s < .01$. However, the difference between the two problem types was greater for users (4.04 vs. 19.28 s) than non-users (11.46 vs. 20.17 s). Infrequent users' latencies were the same for inversion and standard problems, $\underline{F}(1, 22) < 1$. Although we would expect that non-users should not be faster on inversion problems than standard, many students may have used an inversion based strategy that they did not mention, and this decreased their response latencies. It is clear though, that students who said they used inversion were faster on inversion problems than standard problems. Because of the shortcut they used, frequent users should also have faster latencies on inversion problems than the other students. Even though all three groups were equally fast on standard problems, $\underline{F}s(1, 22) < 2.43$, $\underline{p}s > .10$, on inversion problems frequent users were faster than infrequent users and non-users, $\underline{F}s(1, 22) > 5.31$, $\underline{p}s < .05$. Even though inversion users' performance on inversion problems is better than that of other students, it appears this difference is not due to superior skill in calculation.

Because of the computational demands of large problems, infrequent users and non-users should be faster on small problems compared to large problems (i.e. the problem-size effect). No differences between the two problem sizes would be expected for students who frequently used an inversion-based shortcut. Furthermore, students who reported frequent use of shortcuts should be faster on large problems than students who reported infrequent or no ortcut use. Latencies for the three groups did not differ significantly for small problems, $Fs(1, 44) < 1$ (see Table 9). On large problems, however, frequent users were faster than non-users, $F(1, 44) = 9.49$, $p < .01$. The mean latency of infrequent users was intermediate, and did not differ significantly from the means of the other two groups ($ps > .10$). As expected, frequent users were just as fast on large versus as small problems, $F(1, 22) < 1$, and infrequent users and non-users were faster on small problems, $Fs(1, 22) > 9.81$, $ps < .01$. Consistent with the analysis of the accuracy data, these results indicate that the problem-size effect is evident in the performance of non-users and infrequent users, and not evident for frequent users.

Standard problems. On standard problems, some students employed a strategy in which they would subtract the last two numbers first, and then add the first number. For example, on 2

+ 7 - 3, students would subtract 3 from 7 and add the result to 2. This strategy represents a shortcut that was used on standard problems and it is important to examine under what conditions students employed it. Ten students in Grade 2 used subtraction-first at least once, and no correct answers were observed when the difference between the last two numbers was a negative number. Seventeen students in Grade 6 used subtraction-first at least once, and the majority of incorrect answers (12 out of 14) also occurred when the difference between the last two numbers was negative. Although students in both grades used the strategy, it was used correctly only when the difference between the last two numbers was positive.

The proportion of standard problems on which students employed the subtraction-first shortcut was calculated and subjected to a 2(Grade) X 2(Sex) X 2(Task Order) X 2(Problem Size) analysis of variance with repeated measures on the last variable. Students used this strategy on large problems more than on small problems (21.3% vs. 12.8%), $F(1, 72) = 14.44$, $p < .01$. Presumably on larger problems, where students have a harder time getting the answer by simple addition and subtraction, they are more likely to employ an alternative, easier strategy.

Inversion and subtraction-first procedures. If students who frequently use an inversion-based shortcut on inversion problems also use a subtraction-first shortcut on standard problems, we may conclude that the use of shortcuts represents a general strategy some students employ for solving problems. The correlation coefficient between the proportion of inversion use and the proportion of subtraction-first use was .62 for Grade 2 and .57 for Grade 6 ($ps < .01$). Thus students who used the inversion shortcut were more likely to use the subtraction-first shortcut. To determine whether the use of these two shortcuts is related to proficiency in calculation, correlation coefficients were computed between accuracy on standard problems and the frequency of using each shortcut, for each grade. These correlations were negligible, ranging from -.14 to .04. These results imply that individual differences in use of shortcuts are not related to proficiency in calculation (as measured by accuracy on standard problems).

Comprehension Task

Whereas the problem-solving task was used to assess students' spontaneous use of an inversion-based strategy, the comprehension task was used to assess students' ability to identify the appropriate use of an inversion-based strategy (the evaluation component) and explain why that strategy would work (the justification component). Students' responses on the comprehension task also were categorized. A description of all the strategies observed and procedures for scoring are found in Appendix E.

For both lists of inversion problems, responses were categorized as identifying the appropriate use of inversion if the student explained the shortcut before any prompt was given (Yes/No prompt), agreed the shortcut would work (Unconditional Yes), or first checked and then agreed the shortcut would work (Conditional Yes). Although placement of the latter group of responses into this category might be questioned, a majority of students who gave a conditional yes response for one list (large or small) gave an unconditional yes response for the other list (6 out of 10). Furthermore, all of the students who gave a conditional yes response were able to justify inversion later on. Responses were also classified in the justification component. If the student indicated that the last two numbers cancel each other and therefore the answer was the first number, his or her response was coded as justifying inversion. The standard problems were only included in the task to determine whether students would apply the principle of inversion to problems where it would not work. Only two students in Grade 2 incorrectly applied the principle of inversion to standard problems. Responses of these students' for the inversion problems were not coded as identifying the appropriate use of inversion or justifying inversion. Other than in the calculation of reliability, standard problems were not included in the subsequent analyses.

An independent rater, scored responses from 40% of all the subjects, and reliability with the experimenters scoring was calculated for each combination of problem size and problem type. For the evaluation component, reliabilities for the small inversion, large inversion, and standard problems were .94, .94, and 1.00 respectively. For the justification component, reliabilities for the small inversion, large inversion, and standard problems were .75, .84, and .75 respectively. Reliabilities were lower for the justification component because of the complexity of the responses students gave. For example responses often included two or three sentences that were ambiguous. Reliabilities for only those responses that the experimenter coded as inversion-based on were .87 for small problems and .96 for large problems. Because analyses of the justification component employ only the inversion-based responses on inversion problems, it is appropriate to use the second set of reliabilities (.87 and .96).

Evaluation Component

Classification of responses for both large and small problems revealed two groups of students: (a) students who identified the appropriate use of inversion for both problem sizes; and (b) students who did not identify the appropriate use of inversion for both problem sizes. The distribution of students who could and could not identify inversion on large and small problems was examined across several variables. Some results are clear from examining the percentage of students who recognized inversion for each of the variables. Similar numbers of boys as girls (77% vs. 76%) and students who started with the problem-solving task as those who started with the comprehension task (78% vs. 74%) identified appropriate inversion use. As well, students recognized inversion equally often on small problems as large problems (76%). Grade 6 students recognized inversion more frequently than Grade 2 students (90% vs. 62.5%), X^2 (1, $\underline{N}$ = 80) = 8.88, $\underline{p}$s < .01. Although the difference between Grade 6 and Grade 2 students' use of inversion strategies in problem solving was unexpectedly small, students in Grade 6 were able to correctly evaluate an inversion procedure much more frequently than students in Grade 2.

Justification Component

For the justification component, the distribution of students who could and could not explain the principle of inversion was examined in relation to several variables. Chi-squares were calculated to compare whether or not justification of inversion was independent of sex, task order, problem size, and grade. For both small and large problems, justification was independent of sex (66.6% for boys vs. 78.0% for girls on small problems and 71.8% for boys vs. 83.0% for girls on large problems) and task order (73.1% for problem-solving first vs.

71.8% for comprehension first on small problems and 82.9% for problem-solving first vs. 71.8% for comprehension first on large problems), X^2 (1, *N* = 80) < 1.42, *ps* > .10. As well, justification was independent of problem size (72.5% for small vs. 77.5% for large), X^2 (1, *N* = 160) = .53, *p* > .50. However, justification varied as a function of grade for both problem sizes, X^2 (1, *N* = 80) > 4.58, *ps* < .05. Grade 6 students were able to justify inversion more frequently than Grade 2 students (87.5% vs. 57.5% on small and 87.5% vs. 67.5% on large). Although Grade 6 students did not use inversion-based strategies spontaneously in problem solving much more frequently than Grade 2 students, they were able to justify inversion more frequently.

Between-Task Analysis

Response Patterns

To examine relations among the tasks and make inferences about possible forms of understanding associated with them, the performance of each subject was categorized as successfully or unsuccessfully performing each task. Flavell and Wohlwill (1969) argued that an important distinction needs to be made when deciding the criteria for determining success on a task. According to Flavell and Wohlwill, competence refers to the formal representation of some operation, and performance refers to whether the given operation, if functional, will be called into play. Using this distinction, success can be evaluated for each task in two ways. Based on a first-in-competence criterion, students would have to show an inversion-based response at least once, indicating that the necessary knowledge is available but it is not used consistently in performance. Based on an always-in-performance criterion, students would have to demonstrate accurate use of inversion-based knowledge on the majority of trials in a task.

Because the performance required to meet the first-in-competence criterion is different from that of the always-in-performance criterion, it was desirable to categorize each student's performance based on both criteria. Students were rated on each of the three activities (problem-solving task, evaluation component, and justification component) and were categorized on the basis of their ability to successfully apply the principle of inversion. For example, a student may have spontaneously used a shortcut based on the principle to solve inversion problems in the problem-solving task, but not recognized all the instances of inversion in the evaluation component or explained the principle in the justification component. Using the always-in-performance criterion the student would be categorized as successfully performing the problem-solving task, and unsuccessfully performing the evaluation component and the justification component.

After categorization, patterns across the activities were examined. If understanding of inversion is a all or none phenomenon, it would be expected that a student would be successful on all activities, or unsuccessful on all the activities. Based on the first-in-competence criterion, 47.5% of students successfully performed all the tasks, 1.25% unsuccessfully performed all the tasks, and 51.25% successfully performed some but not all the tasks. Based on the always in performance criteria, 32.5% of students successfully performed all the tasks, 13.75% unsuccessfully performed all tasks, and 53.75% successfully performed some but not all the tasks. These results are consistent with the conclusion that different forms of understanding co-exist.

Relations Among Activities

Guttman Scaling was used to examine the possible developmental sequence in the acquisition of different forms of understanding (Guttman, 1944). Wohlwill (1973) argued that despite the original role of Guttman Scaling in the study of presumably continuous dimensions of attitude, matrices of response patterns at the core of the model suggest the presence of a set of discrete responses, such as encountered in the study of developmental sequences. Patterns of success across the three activities were analyzed to determine whether they formed a Guttman scale (Torgerson, 1958). The activities were ordered from hardest to easiest based on the number students that were successful on each activity. The degree to which the data fit this pattern was calculated. Specifically, each deviation from the expected pattern was counted as an error, these errors were accumulated, and two standardized coefficients were computed to determine whether the activities conformed to a Guttman scale. The coefficient of reproducibility was calculated to determine how well the data fit that model, and the coefficient of scalability was calculated to determine whether the fit to the model was due to chance. Guttman (1944) suggested that for a pattern of responses to be considered a scale, the coefficient of reproducibility should at least .90 and the coefficient of scalability should be at least .60.

In the evaluation and justification components students were required to identify and explain an abstract principle about mathematics. In the problem-solving task students were required to use component procedures based on that principle. If abstract principles are acquired before component procedures, then evaluation or justification should have been least difficult, and problem-solving should have been most difficult.

Guttman analysis revealed that, when using the always-in-performance criteria, problem-solving was the hardest activity, followed by justification, and evaluation (see Table 10). The data fit this model very well, with the coefficient of reproducibility being .97 and the coefficient of scalability being .88. When the subjects were subdivided into groups by grade and task order, the data still fit the model very well (see Table 11). When using the first-in-competence criteria, the model was the same as above but the data did not fit as well (see Table 10). The coefficient of reproducibility was .91 and the coefficient of scalability was .58, although these numbers still indicate that the data fit the model. In general, data for Grade 6 students fit the model better than data for students in Grade 2. The influence of task order needs more detailed attention. Performance of students who received the comprehension task first did not fit the model as well because their experience with inversion on this task increased their performance on the problem-solving task. All of the scales produced, however, are consistent with the conclusion that an understanding of principles develops first, followed by an understanding of component procedures.

GENERAL DISCUSSION

The present experiment was designed to examine (a) the conditions that influence students' performance on three activities designed to assess understanding of inversion, (b) whether different forms of understanding are associated with abstract principles and component procedures, and (c) whether a developmental sequence exists in the acquisition of these different forms of understanding.

Analyzing students' performance on three different activities allows for observation of many behaviors related to understanding, and also for examination of conditions that influence those behaviors. If different behaviors are observed for different activities, and different conditions influence performance on each activity, we may conclude that understanding is represented in terms of different forms. In the problem-solving task some students in both grades used solution procedures consistent with the principle of inversion. These students were more accurate and had faster latencies on inversion problems than other students. Although students who used procedures based on inversion were more proficient on inversion problems, they were not more proficient in mathematics generally. Because students who used a shortcut on inversion problems were more likely to use a shortcut on standard problems, we can hypothesize that these students were more flexible in their thinking than other students. Somewhat surprisingly, given the amount of arithmetic instruction in school, frequency of use increased only slightly between Grade 2 and Grade 6 and there was considerable variability in both grades. The use of inversion-based procedures was also influenced by a student's prior experience on a task in which with the principle of inversion could be involved. Specifically, students who were asked to identify and justify the inversion principle were much more

likely than other students to use procedures based on i
spontaneously. This facilitative effect was the same
grades.

Although many students in both grades identified t
appropriate use of a strategy based on inversion, Grade
students did so more frequently than students in Grade
Students in Grade 6 were also able to justify inversion
frequently than students in Grade 2. Previous experien
solving inversion problems did not influence a student'
performance on either the evaluation component or the
justification component. If understanding were uniform
would expect that the conditions influencing performanc
the same for all of the tasks. Although there was no g
effect on the problem-solving task, there were large gr
effects for the comprehension tasks. What might accoun
difference? The spontaneous use of inversion-based sho
not have increased significantly between Grades 2 and 6
on the problem-solving task students were required to g
the strategy on their own. Strategy generation may be
difficult for all ages of students. In the comprehensi
students were required to recognize and justify a strat
has already been described to them. Because of their e
at school, students in Grade 6 may have been much more
these types of activities than students in Grade 2. Fu

previous experience influenced the spontaneous use of procedures based on inversion but did not influence evaluation or justification of inversion. These results are consistent with the conclusion that understanding is not uniform, but rather it is differentiated into many forms.

Cooney and Ladd (1992) argued that children's reports of strategy use may not be veridical. Because the data collected in the problem-solving task were based upon students self-reports of strategy use, this issue must be addressed. By classifying subjects into groups on the basis of their self-reports, I could examine accuracy and latency for students who claimed to use shortcuts in varying degrees. If self-reports are not veridical, as Cooney and Ladd suggested, then there should be no relation between students accuracy, latency, and self-reports. If reports are veridical however, students who report frequent use of shortcut should have higher accuracy rates and faster latencies than students who did not report frequent use of shortcuts. Students who reported using a shortcut on inversion problems had higher accuracy rates and faster latencies than other students. These results suggest that the self-reports were veridical.

Examination of students' response patterns across the three activities is important because it takes into account individual differences in student's performance. Results from these analyses support the conclusion that it is best to characterize understanding in terms of different forms, rather than as single construct. For inversion there appears to be a form of understanding associated with the application of procedures based on the principle of inversion, a form associated with the evaluation of procedures to determine their appropriateness, and a form associated with the justification of procedures. Further, these forms of understanding can be developmentally ordered. Analyses support the conclusion that evaluation and justification precedes spontaneous application in the course of acquisition. Consistent with this conclusion is the idea that knowledge in this domain is represented initially in terms of general, relatively abstract principles, which in turn can be used to guide or constrain the constructions of more specific procedures (Bisanz & LeFevre, 1992).

Although results from the present experiment support the conclusion that abstract principles guide the use of component procedures, a few questions must still be addressed. First, even though acquisition of knowledge about inversion can be represented in terms of abstract principles guiding component procedures, acquisition of knowledge in other domains may not occur this way. In future studies it would be important to examine performance in a number of knowledge domains to determine whether the sequence of development in those domains paralleled each other or whether development of understanding was different for different domains.

Second, the mechanisms by which understanding develops must also be examined. Although using Guttman Scaling has helped to identify a developmental sequence in which evaluation precedes justification, which in turn precedes application of procedures, this analysis does not address the question of how understanding develops from one form to another. How does a student who is able to explain why inversion works reach the point where she is able to use an inversion-based shortcut spontaneously to solve problems? To gain increased knowledge about the development of understanding, it is not enough to know which form precedes another in acquisition. Rather, the nature of change between the forms must be examined. One way to examine change would be to use the microgenetic method proposed by Siegler and Crowly (1991). This method involves (a) observations of individual students throughout the period of change, (b) a high density of observations relative to the rate of change within that period, and (c) intensive trial-by-trial analyses intended to infer the processes that gave rise to the change. To examine different forms of understand of inversion one should identify students at different stages, and examine their performance over a number of weeks.

A microgenetic study may also help to address a related question: How do abstract principles influence the construction or selection and subsequent implementation of task-specific procedure (Bisanz & LeFevre, 1992)? Bisanz and LeFevre (1990) propose a possible model of how knowledge about the principle of inversion might enable students to use shortcuts on inversion problems. Abstract principles could be represented in terms of productions (condition-action statements). In a production system these conceptual productions could be modified by knowledge-acquisition productions to create task-specific procedures that enable a student to use shortcuts on certain problems. Use of a microgenetic model would be helpful in (a) identifying task-specific productions and (b) examining students discrete behaviors to make inferences about possible conceptual and knowledge-acquisition productions.

Using a problem-solving activity, an evaluation activity, and a justification activity, I identified different forms of understanding for inversion and the order in which these forms may be acquired. For psychologists interested in providing a full account of remembering and problem solving, an integrated and detailed description of the relation between computational processes and the common notion of understanding is needed (Bisanz & LeFevre, 1992). To facilitate this goal, future research needs to be focused more specifically on the computational processess related to different forms of understanding. Further, research needs to be focused on how these different forms of understanding are represented in different knowledge domains. As well, educators and researcher need to develop instructional methods and assessment techniques that emphasize the diverse nature of understanding.

References

Alberta Education, Student Evaluation Branch. (1990). Diagnostic mathematics program.

Anderson, J. R., Boyle, C., & Yost, G. (1983). The geometry tutor. In Proceedings of the international joint conference on artificial intelligence (pp 1-7). Los Angeles, CA: International Joint Conference on Artificial Intelligence.

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

Baroody, A. J., & Ginsburg, H. P. (1986). The relationship between initial meaningful and mechanical knowledge of arithmetic. In J. Hiebert (Ed.), Conceptual and procedural knowledge: The case of mathematics (pp. 75-112). Hillsdale, NJ: Erlbaum.

Bisanz, J., & LeFevre, J. (1987, June). Development of heuristic use in mental arithmetic. Poster presented at the annual meeting of the Canadian Psychological Association, Vancouver, BC.

Bisanz, J., & LeFevre, J. (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. (1992). Understanding elementary mathematics. In J. Campbell (Ed.), The nature and origins of mathematical skills (pp. 113-136). Elsevier Science Publishers.

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

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.

Byrnes, J.P., & Wasik, B.A. (1991). Role of conceptual knowledge in mathematical procedural learning. Developmental Psychology, 27, 777-786.

Campbell, J. I. D. (1987). Network interference and mental multiplication. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 109-123.

Campbell, J. I. D., & Clark, J. M. (1989). Time course of error priming in number-fact retrieval: Evidence for excitatory and inhibitory mechanisms. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 920-929.

Case, R. (1985). Intellectual development: Birth to adulthood. Orlando: Academic Press.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 1, pp. 7-75).

Charles, R., & Lester, F. (1982). Teaching problem solving: What why and how. Palo Alto, CA: Dale Seymour Publications.

Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatley, G., Trigatti, B., & Perlwitz, M. (1991). Assessment of a problem-centerd second-grade mathematics project. Journal for Research in Mathematics Education, 22, 3-29.

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.

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

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new era of cognitive-developmental inquire. American Psychologist, 34, 906-911.

Flavell, J. H., & Wohlwill, J. F. (1969). Formal and functional aspects of cognitive development. In D. Elkind & J. H. Flavell (Eds.), Studies in cognitive development (pp. 67-120). New York: Oxford University Press.

Fuson, K. C., & Hall, J. W. (1983). The acquisition of early number word meanings. In H. Ginsburg (Ed.), The development of children's mathematical thinking (pp. 49-107). New York:

Academic Press.

Gelman, R., & Gallistel, C. R. (1978). The child's understanding of number. Cambridge, MA: Harvard University Press.

Gelman, R., & Meck, E. (1983). Preschoolers' counting: Principles before skill. Cognition, 13, 343-359.

Greeno, J. G. (1983). Forms of understanding in mathematical problem solving. In S. G. Paris, G. M. Olson, & H. W. Stevenson (Eds.), Learning and motivation in the classroom (pp. 83-111). Hillsdale, NJ: Lawrence Erlbaum Associates.

Greeno, J. G., & Riley, M. S. (1987). Processes and development of understanding. In F. E. Weinert & R. H. Kluwe (Eds.), Metacognition, motivation, and understanding (pp. 289-313). Hillsdale, NJ: Lawrence Erlbaum Associates.

Guttman, L. (1944). A basis for scaling qualitative data. American Sociological Review, 9, 139-150.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), Conceptual and procedural knowledge: The case of mathematics (pp. 1-27). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Krulik, S., & Rudnick, J. (1988). Problem solving: A handbook for elementary school teachers. Toronto: Allyn and Bacon, Inc.

Krzanowska, E. (1988). Relation between procedural skill and understanding in multiplication. Unpublished Master's thesis, University of Alberta, Edmonton, Alberta, Canada.

Mayer, R. E. (1985). Structural analysis of science prose: Can we increase problem-solving performance? In B. K. Britton & J. B. Black (Eds.), Understanding expository text: A theoretical and practical handbook for analyzing explanatory text (pp. xxx-xxx). Hillsdale, NJ: Erlbaum.

Putnam, R. T., deBettencourt, L. U., & Leinhardt, G. (1990). Understanding of derived-fact strategies in addition and subtraction. Cognition and Instruction, 7, 245-285.

Schoenfeld, A. H. (1985). Mathematical problem solving. Toronto: Academic Press, Inc.

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

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

Silver, E. A. (1986). Using conceptual and procedural knowledge: A focus on relationships. In J. Hiebert (Ed.). Conceptual and procedural knowledge: The case of mathematics (pp. 181-198).

Hillsdale, NJ: Lawrence Erlbaum Associates.

Silver, E. A. (1985). Research on teaching mathematical problem solving: Some underrepresented themes and needed directions. In E.A. Silver (Ed.), Teaching and learning mathematical problem solving: Multiple research perspectives (pp. 247-265).

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.

Torgerson, W.S. (1958). Theory and methods of scaling. New York, Wiley.

Wohlwill, J.F. (1973). The study of behavioral development. New York, Academic Press.

Appendix A

Application of Procedures Problem Set

1)	7 + 2 - 2 =	(small inversion)
2)	2 + 4 - 5 =	(small standard)
3)	3 + 24 - 26 =	(large standard)
4)	9 + 27 - 27 =	(large inversion)
5)	4 + 8 - 6 =	(small standard)
6)	5 + 29 - 27 =	(large standard)
7)	8 + 6 - 6 =	(small inversion)
8)	7 + 22 - 21 =	(large standard)
9)	2 + 25 - 25 =	(large inversion)
10)	6 + 2 - 3 =	(small standard)
11)	8 + 26 - 28 =	(large standard)
12)	5 + 7 - 7 =	(small inversion)
13)	6 + 22 -22 =	(large inversion)
14)	9 + 7 - 5 =	(small standard)
15)	4 + 28 - 28 =	(large inversion)
16)	3 + 4 - 4 =	(small inversion)

Appendix B

Comprehension Task Problem Sets

Small Inversion List

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 Inversion List

5 + 22 - 22 =

8 + 27 - 27 =

4 + 28 - 28 =

7 + 25 - 25 =

3 + 29 - 29 =

6 + 23 - 23 =

2 + 26 - 26 =

9 + 24 - 24 =

Standard List

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 C

Briefing and Instructions

A) First Session: "We are trying to find out what stud[redacted] about different math problems. Can you help me with thi (students name)? I will give you some problems to think and then I will ask you some questions. Some of the pro will be easy and some will be hard. What I am interest how you think about math, so don't worry if you have tr[redacted] some of the problems, this is not a test. All I want y[redacted] is try your best. The answers that you give will be rec[redacted] the video camera, and I will also write your answers do[redacted] this will help me to remember your answers later on. S[redacted] begin?"

B) Problem-solving task Instructions: "Today I would li[redacted] solve a number of addition and subtraction problems and how you solved them. I will show you a set of problems time and when you get an answer, say it out loud. I wi[redacted] ask you to tell me what you were thinking as you solved problem. To give you some practice, I want you to tell you think the answer is for the following problems." P[redacted] two-term problems. "Now I'm going to give you some diff[redacted] problems, and I want you to tell me the answers, and h[redacted] the answers."

"So those were some practice problems. Now lets continue with new problems. Remember, I want you to tell me the answer first, and then tell me what you were thinking as you solved the problem. Are you ready to begin?"

C) Comprehension Task

Instructions: "Now I want you to look at different lists of problems, and for each list, I want you to tell me what is the same about all the problems in the list. Here is the first list. Look at all the problems carefully and tell me how they are all alike."

Prompt: If the child doesn't know how they are alike, show him or her how they are alike. For example: "For all these problems you add one number, and then take away that same number. Does that work down here? Is that the same for all these problems?" Then ask the child: "If you had to tell another student how these problems are alike, what would you tell them?" When satisfied that the child understands how the problems are similar continue with the task.

<u>Comprehension Task Text</u>

For each inversion set (a+b-b) say: "A boy/girl I know says that if you start with a certain number (point at the "a" numbers) and you add 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. He/she says that in any problem with plus a number and minus that same number you really don't need to add and take away those two numbers, the answer will always the number you started with. So here (point to a problem) the answer would be ...(have student respond). Well any way, that's what he/she says. What do you think? Would his/her way of doing it give you the right answer for all these problems, or would you have to add and subtract each number separately (evaluation of procedures component)? Why do you think that (justification of procedures component)?"

For second set of inversion problems say: "Now this list of problems is similar to the last list of problems that we looked at. See you add a number and take away the same number. Is that the same for all the numbers in the list?"

"The same boy/girl says that, like the last list of problems, in any problem with plus a number and minus that same number you really don't need to add and take away those two numbers, the answer will always the number you started with. So here (point to a problem) the answer would be ... Well any way, that's what he/she says. What do you think? Would his/her way of doing it give you the right answer for all these problems, or would you have to add and subtract each number separately? Why do you think that?"

For the standard set say: "A boy/girl I know says that if you start with a certain number (I will point at the "a" number) and you add another number (I will point at the "b" number), and then take away a different number (I will point at the "c" number), the answer is always going to be the first number you started with. He/she says that in any problem with plus a number and minus a different number you really don't have to add and take away those two numbers, the answer will always be the number you started with. So here (point to a problem) the answer would be... Well any way that's what he/she says. What do you think? Would his/her way of doing it give you the right answer for all of these problems, or would you have to add and subtract each number separately? Why do you think that?"

Appendix D

Categories of Strategy Use: Application of Procedures Task

Each response was classified according to (a) the strategy used to solve a problem and (b) the specific steps that the student may have taken to carry out that strategy. Sometimes students used more than one step.

Left to Right: When using this strategy students added and subtracted the numbers moving from the left to the right of the problem. Sometimes a student started addition with the largest number even if it was the second number. This approach was given a different code than if they started with the first number.

Example: 7 + 4 - 9 =

"7 + 4 = 11"

"11 - 9 = 2"

Separation: When using the left to right strategy, some students would separate a larger number into smaller subsets (ones, tens, twenties)

Example: 2 + 22 - 21 =

"2 + 22 = (2 + 2) + 20 = 24"

"24 - 21 = (4 - 1) + (20 - 20) = 3"

Counting: Sometimes a student counted aloud or on fingers, while solving a problem from left to right.

Example: 9 + 2 - 3 =

"count 2 up (9, 10, 11)"

"count 3 down (11, 10, 9, 8)"

Suspected Counting: Sometimes it was not easy to tell if a student was counting or not. Counting was suspected if: (a) the student's mouth or head moved as if they were counting, (b) the student manipulated his or her fingers, and (c) the student's hands appeared to move under the table while he or she was solving a problem, and were above table when he or she was not solving a problem.

Derived Fact: Sometimes a student used some derived fact to help them solve a problem while moving from left to right.

Example: 8 + 4 - 6 =

"(8 + 4 = 12) (know that 6 plus 6 is 12

therefore 12 - 6 must be 6)"

No Visible Procedure: In many cases there was no visible evidence that a student had used some type of intermediate step.

Subtraction-first: Using this strategy, students reduced the problem to a simpler form by subtracting first.

Example: 9 + 28 - 24 =

"(28 - 24 = 4), (4 + 9 = 13)"

Separation: Here the student separated larger numbers into subsets.

Counting: Here the student counted aloud or on fingers.

Derived Fact: Here the student uses some derived fact that they knew to help solve the problem.

Negation: In negation the student would add the first two numbers together, and then would realize that he or she had to subtract the same number that was added. The student would remember the first number and give it as the answer rather than subtracting the last number. It was difficult to distinguish this strategy from both the left-to-right strategy and the inversion strategy. To categorize a response as negation there had to be evidence that the student added the first two numbers (eg. giving the sum of the first two numbers). As well, there had to be evidence that he or she did not subtract the last number (eg. saying "It just goes back to first number.").

Example: 9 + 28 - 28 =

"(9 + 28 = 37) (take away the 28, so it just goes back to 9)"

Inversion: When using this strategy the student stated "the last two numbers are the same" or "the last two numbers cancel out" and quickly stated the first number as the answer. There must have been no indication that the student added or subtracted any of the numbers in the problem

Example: 9 + 28 - 28 = 9

"(28 - 28 = 0) or (28's cancel)" or "(take away the same number so it is the first number)"

Clear Strategy: Here the student used some other clear strategy to solve all problems.

Example: the student always gave the first number as the answer to the problem (position rule, magnitude rule).

Other: Anything that did not fit the above categories was classified as "other."

Appendix E

Comprehension Task: Evaluation Component

Classifying responses in evaluation of procedures component involved coding the first response students made to the following question: "Would her/his way of doing it give you the right answer, or would you have to add and subtract each number separately?"

Yes (no prompt): Here the student said the inversion strategy would work before hearing the prompt. For example, a response was classified as yes, no prompt if when the student was asked to explain how all the problems in the list were the same, they said "the last two equal zero so the answer is the first number," or "the last two cancel out so the answer is the first number."

Unconditional Yes: Here the student agreed that the inversion strategy could be used, and it would be correct. For example, a response was classified as unconditional yes if after hearing the prompt the student said "her way would work," or "yes they are right."

Conditional Yes: Here the student indicated that you could use the inversion strategy but did so based on some condition. For example, the student might have said "yes it could work but you would have to add and subtract to make sure it is right." If there was evidence the student looked through the problems to check if the first number was the answer, and then said "yes her way would be right", their response was classified as conditional yes.

Unconditional No: Here the student indicated that the inversion strategy could not be used. For example, when asked if the girl's way would work the student said "you would have to add and subtract all the numbers" or said "no her way would not work."

Conditional No: Here the student indicated that no the girl's strategy would not work but the student's answer was based on some condition. For example, the student might have said "no I don't think her way would work but I would have to check."

No, no prompt: Here the student said the inversion strategy would not work before he or she heard the prompt. For example, before hearing the prompt the student said "on these ones her way would not work."

Maybe: Here the student thinks the inversion strategy would work but he or she was not sure.

I don't know: The student did not know one way or another if the inversion strategy would work.

Comprehension Task: Justification Component

Classifying responses in the justification component involved analyzing the justifications students gave for their responses in the evaluation of procedures component. This included all responses after the question "Why do you think that?" was asked.

<u>Add and Subtract</u>: The student indicated that you must add and subtract all the numbers to get the right answer. They did not attempt to solve any of the problems. For example, students said "you must add and subtract to make sure the answers are right", or "because that's the right way to do it", or "the numbers are different in a problem, therefore you must add and subtract the numbers", or "the last two numbers don't equal zero therefore you have to add and subtract."

<u>Empirical Proof</u>: Some students gave an empirical proof to a specific problem to justify their answer. For example, the student added and subtracted all the numbers in the problem to see if saying the first number is the answer was right. There must have been evidence that they added and subtracted all the numbers. The student might have said, "his way of doing it will give you the right answers see, 4 + 8 = 12, and 12 - 8 = 4" (for problem 4 + 8 - 8 =).

<u>I don't know</u>: The student could not give a justification for his or her answer.

<u>Other</u>: The student gave some clear, consistent justification that did not fit in the other categories. For example, the student said "the first number is always the answer," or "the smallest number is the answer."

<u>Ambiguous</u>: Here the student gave a reason that was not clear. For example, the student said "because the last two numbers are

higher the 20."

Inversion with an Example: Some students used a specific example to justify inversion but could not state a general rule. There was also no evidence they added all the numbers together to check. For example, the student said "if you add 25 and take away the same 25 you added, you only have seven left." Although he or she gave an explanation for inversion, the student tied it to a specific example. The student could not generalize to a general rule.

Inversion/ General Rule: Here the student used a general rule to explain inversion. Although originally, he or she may have used a specific example, if at any time the student gave a general rule, he or she was placed in this category. For example, the student said "the last two numbers are the same, therefore the answer is the first number", or "you add and subtract the same number, therefore the answer is the first number", or "the last two numbers equal zero, therefore the answer is the first number."

Table 1

Mean Initial Accuracy as a Function of Grade and Problem Size

	Problem Size	
Grade	Small	Large
2	.74	.51
6	.95	.86

Table 2

Mean Initial Accuracy as a Function of Problem Type and Problem Size

	Problem Size	
Problem Type	*Small*	*Large*
Standard	.77	.52
Inversion	.92	.85

Table 3

Mean Initial Accuracy as a Function of Grade, Task Or

Problem Type

	Grade 2	
	Problem Solving First	Comprehens
Problem Type		
Standard	.52	.3
Inversion	.72	.8

	Grade 6	
	Problem Solving First	Comprehens
Problem Type		
Standard	.83	.8
Inversion	.93	.9

Table 3

Mean Initial Accuracy as a Function of Grade, Task Order, and Problem Type

	Grade 2	
	Problem Solving First	Comprehension First
Problem Type		
Standard	.52	.38
Inversion	.72	.89
	Grade 6	
	Problem Solving First	Comprehension First
Problem Type		
Standard	.83	.86
Inversion	.93	.99

Table 4

Percent of Strategy Use: Inversion Problems

	Problem Size	
Strategy	Small	Large
Left to Right	46.8	38.7
Separation	0.0	1.6
Counting	4.0	5.3
Suspected Counting	0.3	0.3
Derived Fact	0.9	0.3
No visible procedure	4[illegible].3	30.3
Other	0.3	0.9
Negation	6.6	6.9
Inversion	43.4	49.4
Clear Strategy	3.1	4.1
Other	0.0	0.9

Table 5

Percent of Strategy Use: Standard Problems

	Problem Size	
Strategy	Small	Large
Left to Right	81.5	69.5
Separation	0.0	5.1
Counting	8.5	10.7
Suspected Counting	0.9	0.9
Derived Fact	6.2	1.5
No visible procedure	65.0	48.4
Other	0.9	2.9
Subtraction-first	14.0	21.2
Separation	0.0	2.5
Counting	0.9	0.0
Suspected Counting	0.9	0.0
No visible procedure	12.2	18.4
Other	0.0	0.3
Clear Strategy	5.3	7.5
Other	0.9	1.7

Table 6

Mean Accuracy as a Function of Grade, Inversion Use, and Problem Type

	Grade 2	
	Standard Problems	Inversion Problems
Inversion Use		
Frequent Users	.41	.99
Infrequent Users	.53	.72
Non-users	.44	.69

	Grade 6	
	Standard Problems	Inversion Problems
Inversion Use		
Frequent Users	.83	.98
Infrequent Users	.82	.93
Non-users	.88	.95

Table 7

Mean Accuracy of a Function of Inversion Use and Problem Size

	Problem Size	
Inversion Use	*Small*	*Large*
Frequent Users	.84	.77
Infrequent Users	.87	.63
Non-users	.84	.64

Table 8

Average Median Latencies (in seconds) as a Function of Inversion Use and Problem Type for Students in Grade 2

	Problem Type	
Inversion Use	Standard	Inversion
Frequent Users	19.28	4.04
Infrequent Users	15.14	12.94
Non-users	20.17	11.46

Table 9

Average Median Latencies (in seconds) as a Function of Inversion Use and Problem Size for Grade 2 Students

	Problem Size	
Inversion Use	Small	Large
Frequent Users	10.51	12.81
Infrequent Users	9.17	18.61
Non-users	8.05	23.58

Table 10

Guttman Scaling: Response Patterns

Always-in-Performance Criterion

Problem -solving	Justification	Evaluation	Number of Subjects
+	+	+	26
+	+	-	4
+	-	+	3
+	-	-	1
-	+	+	22
-	+	-	3
-	-	+	10
-	-	-	11

(continued)

First-in-competence Criterion

Problem -solving	Justification	Evaluation	Number of Subjects
+	+	+	42
+	+	-	0
+	-	+	3
+	-	-	6
-	+	+	15
-	+	-	4
-	-	+	9
-	-	-	1

Note. + indicates success based on the criterion, and - indicates failure.

Table 11

Guttman Scaling Coefficients

	Always-in-Performance Criterion	
	Reproducibility	Scalability
Grouping		
All students (both orders)	.97	.88
Task Order 1	1.00	1.00
Task Order 2	.93	.76
Grade 2 (both orders)	.95	.84
Task Order 1	1.00	1.00
Task Order 2	.89	.72
Grade 6 (both orders)	.98	.93
Task Order 1	1.00	1.00
Task Order 2	.97	.80

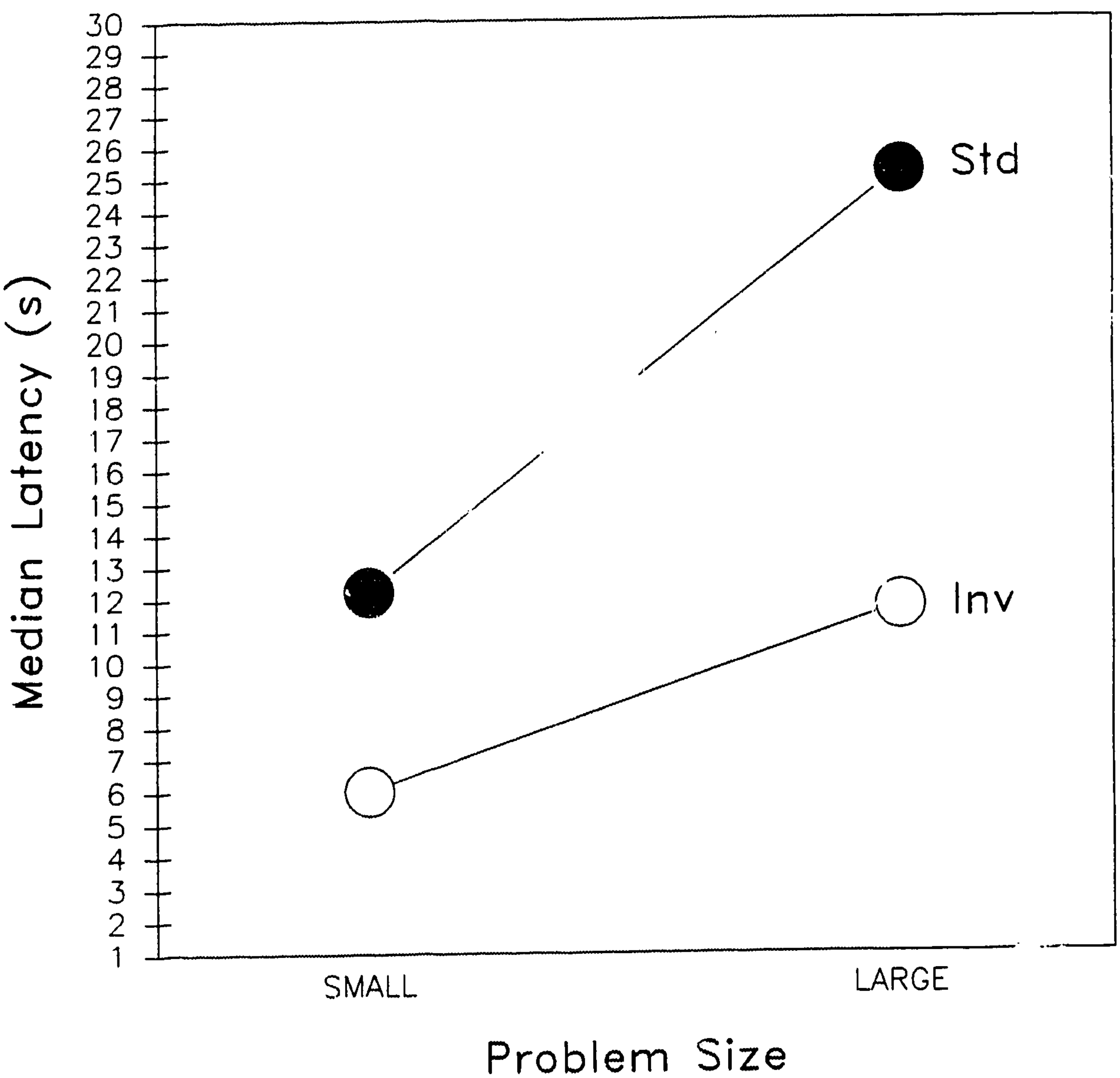
(continued)

	First-in-Competence Criterion	
	Reproducibility	Scalability
<u>Grouping</u>		
All Students (both orders)	.91	.58
Task Order 1	.95	.78
Task Order 2	.90	.52
Grade 2 (both orders)	.88	.61
Task Order 1	.94	.75
Task Order 2	.91	.70
Grade 6 (both orders)	.97	.73
Task Order 1	.97	.71
Task Order 2	.97	.75

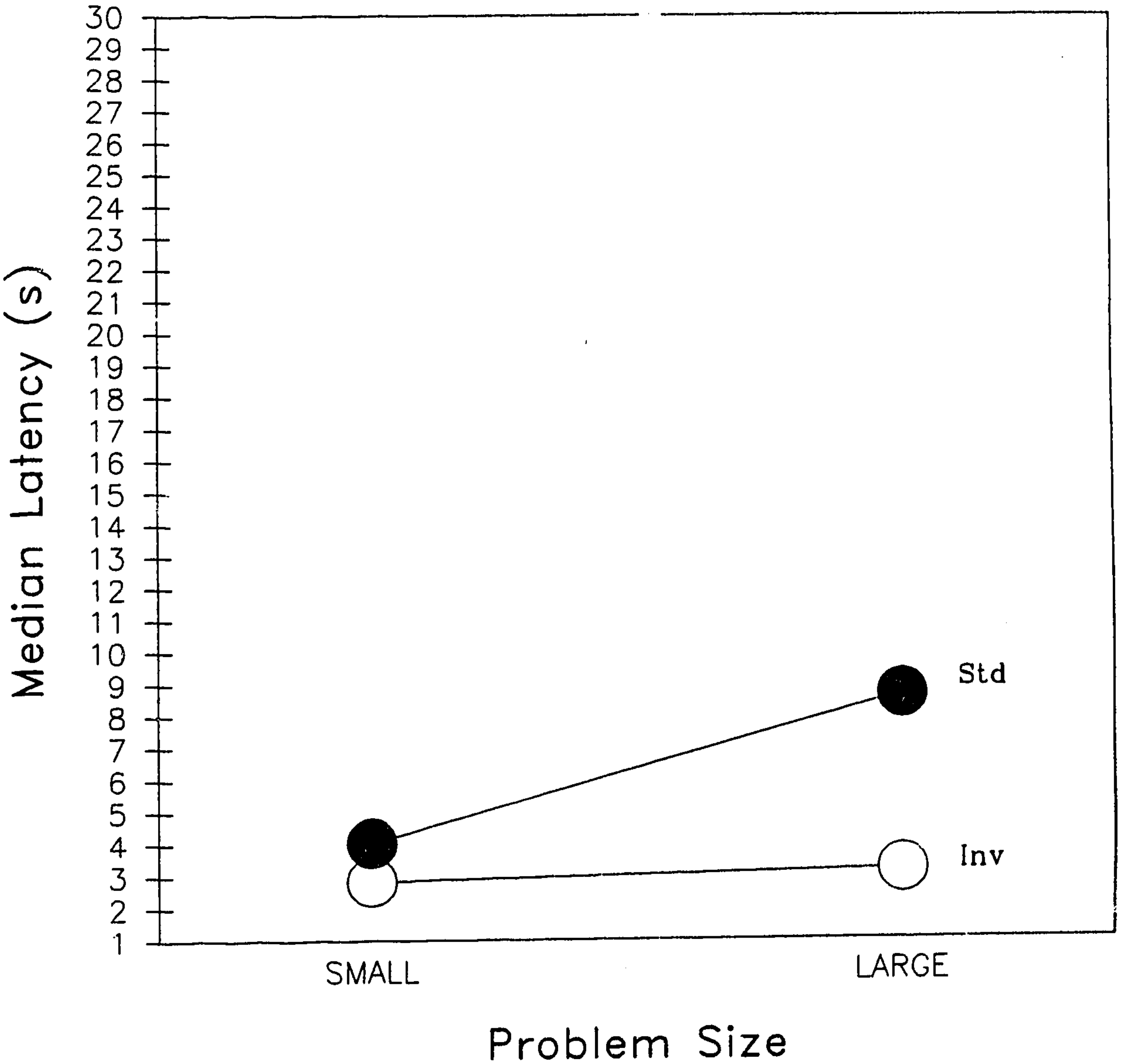
Figure Captions

Figure 1. Latencies as a function of problem size and problem type for Grade 2 students.

Figure 2. Latencies as a function of problem size and problem type for Grade 6 students.



Median Latency (s)
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
Std
Inv
SMALL
LARGE
Problem Size


Median Latency (s)
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
Std
Inv
SMALL
LARGE
Problem Size