Journal of Educational Psychology
1999, Vol. 91, No. 1, 175–189


0022-0663/99/$3.00

# Conceptual and Procedural Knowledge of Mathematics: Does One Lead to the Other?

Bethany Rittle-Johnson and Martha Wagner Alibali
Carnegie Mellon University

This study examined relations between children's conceptual understanding of mathematical equivalence and their procedures for solving equivalence problems (e.g., 3 + 4 + 5 = 3 + __). Students in 4th and 5th grades completed assessments of their conceptual and procedural knowledge of equivalence, both before and after a brief lesson. The instruction focused either on the concept of equivalence or on a correct procedure for solving equivalence problems. Conceptual instruction led to increased conceptual understanding and to generation and transfer of a correct procedure. Procedural instruction led to increased conceptual understanding and to adoption, but only limited transfer, of the instructed procedure. These findings highlight the causal relations between conceptual and procedural knowledge and suggest that conceptual knowledge may have a greater influence on procedural knowledge than the reverse.

In many domains, children must learn both fundamental concepts and correct procedures for solving problems. For example, mathematical competence rests on children developing and connecting their knowledge of concepts and procedures. However, the developmental relations between conceptual and procedural knowledge are not well-understood (Hiebert & Wearne, 1986; Rittle-Johnson & Siegler, in press). Delineating how the two forms of knowledge interact is fundamental to understanding how learning occurs.

This issue is of practical, as well as theoretical, importance. The widespread observation that many children perform poorly in school mathematics highlights the need for improved instruction. Educators and policy makers are placing increased emphasis on teaching the conceptual basis for problem-solving procedures (National Council of Teachers of Mathematics [NCTM], 1989), in hopes that increasing conceptual understanding will lead to improved problem-solving performance. To design effective instruction, one must carefully delineate key mathematical concepts and their associated procedures, identify what children at various ages understand and what they struggle to learn, and examine how instruction influences children's acquisition of both concepts and procedures.

The purpose of the present study was to explore the relations between conceptual and procedural knowledge in children learning the principle that the two sides of an equation represent the same quantity. Specifically, the study investigated how instruction about the concept of mathematical equivalence influences children's problem-solving procedures and how instruction about a problem-solving procedure influences children's conceptual understanding of equivalence. In addressing these issues, we also identified what aspects of equivalence fourth- and fifth-grade students understand, what aspects they do not understand but can easily learn, and what aspects they have difficulty learning.

## Impact of Conceptual Knowledge on Procedural Knowledge

We define *conceptual knowledge* as explicit or implicit understanding of the principles that govern a domain and of the interrelations between pieces of knowledge in a domain. We define *procedural knowledge* as action sequences for solving problems. These two types of knowledge lie on a continuum and cannot always be separated; however, the two ends of the continuum represent two different types of knowledge.

These two types of knowledge do not develop independently. Indeed, it is likely that children's conceptual understanding influences the procedures they use. Several theories of knowledge acquisition suggest that procedure generation is based on conceptual understanding (e.g., Gelman & Williams, 1997; Halford, 1993). Children are thought to use their conceptual understanding to constrain procedure discovery and to adapt their existing procedures to novel tasks (e.g., Gelman & Gallistel, 1978; Gelman & Meck, 1986; Siegler & Crowley, 1994).

Four types of evidence from research on mathematics learning support the idea that conceptual understanding

This research was supported by a National Science Foundation Graduate Fellowship to Bethany Rittle-Johnson and by a Faculty Development Award from Carnegie Mellon University to Martha Wagner Alibali. We are grateful to the students, teachers, and administrators at Word of God and St. Maurice schools in the Diocese of Pittsburgh, who made this research possible. We thank Leslie Petasis and Victor Ortega for assistance with coding, and Lisa Haverty, Neil Heffernan, Ken Koedinger, Adisack Nhouyvanisvong, Michelle Perry, Robert Siegler, and Doug Thompson for helpful discussions and constructive feedback. Portions of these results were presented at the 28th Annual Symposium of the Jean Piaget Society, June 1998, Chicago.

Correspondence concerning this article should be addressed to Bethany Rittle-Johnson or Martha Wagner Alibali, Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213-3890. Electronic mail may be sent to Bethany Rittle-Johnson at br2e@andrew.cmu.edu or to Martha Wagner Alibali at alibali@andrew.cmu.edu.

plays a role in generation and adoption of procedures (see Rittle-Johnson & Siegler, in press, for a more detailed review). First, children with greater conceptual understanding tend to have greater procedural skill. For example, children who have a better understanding of place value are more likely to successfully use the borrowing procedure for multidigit subtraction (Cauley, 1988; Hiebert & Wearne, 1996). This association between conceptual and procedural knowledge has also been found in many other domains within mathematics, including counting (Cowan, Dowker, Christakis, & Bailey, 1996), single-digit arithmetic (Baroody & Gannon, 1984; Cowan & Renton, 1996), fraction arithmetic (Byrnes & Wasik, 1991), and proportional reasoning (Dixon & Moore, 1996). Note that these studies demonstrate that conceptual and procedural knowledge are related, but they do not show whether the two types of knowledge influence one another.

Second, in several domains, conceptual understanding precedes procedural skill. Indeed, there is some (albeit limited) evidence that level of conceptual understanding predicts future procedural knowledge. For example, some evidence suggests that preschoolers understand principles of counting when they first learn to count (Gelman & Meck, 1983, 1986). Similarly, in single-digit addition, kindergartners understand the principle of commutativity for addition (i.e., the idea that changing the order of the addends does not change the sum), and can recognize the validity of a procedure based on it, before they use this more advanced procedure themselves (Baroody & Gannon, 1984; Cowan & Renton, 1996; Siegler & Crowley, 1994). Conceptual knowledge also seems to precede procedural knowledge in several other mathematical domains, including integer addition and subtraction (Byrnes, 1992), fraction addition (Byrnes & Wasik, 1991), and proportional reasoning (Dixon & Moore, 1996). In multidigit arithmetic, conceptual understanding not only precedes the use of correct procedures for many children, but also predicts future procedural skill (Hiebert & Wearne, 1996). These findings suggest that conceptual knowledge has a positive influence on procedural knowledge; however, other factors, such as IQ or motivation, may account for the apparent relation.

Third, instruction about concepts as well as procedures can lead to increased procedural skill. In multidigit arithmetic, several studies have shown that instruction that includes a conceptual rationale for procedures leads to greater procedural skill than conventional, procedure-oriented instruction (Fuson & Briars, 1990; Hiebert & Wearne, 1992, 1996). However, another study found that similar instruction did not improve procedural skill for children who already used an incorrect subtraction procedure (Resnick & Omanson, 1987). Thus, the empirical evidence on this issue is mixed. Further, the results from these studies must be interpreted cautiously because the research did not examine the independent effect of conceptual instruction, and the studies did not include the control groups necessary to draw causal conclusions.

Finally, one study has provided suggestive evidence that increasing conceptual knowledge leads to procedure generation. Perry (1991) found that instruction on the concept of mathematical equivalence led a substantial minority of fourth- and fifth-grade students to generate a correct procedure for solving equivalence problems. This study did not include a no-instruction control group; however, other studies have shown that children rarely generate a correct procedure for solving equivalence problems in the absence of instruction (e.g., Alibali, 1999). Thus, it seems fair to conclude that, in Perry's study, gains in conceptual knowledge led to generation of correct procedures. To date, Perry's work provides the only causal evidence that increasing children's conceptual knowledge of mathematics leads to increased procedural knowledge.

## Impact of Procedural Knowledge on Conceptual Knowledge

Overall, the literature suggests that conceptual understanding plays an important role in procedure adoption and generation. However, it seems likely that this relationship is not a unidirectional one. Instead, conceptual and procedural knowledge may develop iteratively, with gains in one leading to gains in the other, which in turn trigger new gains in the first. Thus, procedural knowledge could also influence conceptual understanding. Indeed, some theories of knowledge acquisition postulate that knowledge begins at an implicit, procedural level and over time becomes increasingly explicit and well-understood (e.g., Karmiloff-Smith, 1986, 1992). However, these theories have not focused on academic domains, and evidence for the impact of procedural knowledge on conceptual knowledge within mathematics is sparse.

Under some circumstances, children first learn a correct procedure and later develop an understanding of the concepts underlying it. For example, several studies have shown that 3- and 4-year-olds can count correctly before they understand certain counting principles, such as the irrelevance of counting order (Briars & Siegler, 1984; Frye, Braisby, Love, Maroudas, & Nicholls, 1989; Fuson, 1988; Wynn, 1990). Because children do not receive instruction on counting principles, it seems likely that they abstract the principles from their counting experiences (Siegler, 1991). Similarly, kindergartners (who know how to add) often understand the principle of commutativity for addition, even though they have not yet been instructed on the principle. Understanding of this principle is probably abstracted from their experience solving addition problems (Siegler & Crowley, 1994).

In contrast to these results, some studies have suggested that gains in procedural knowledge do not necessarily lead to increased conceptual knowledge. In domains such as fraction multiplication and multidigit subtraction, many children learn correct procedures but never seem to understand the principles that justify them (e.g., Byrnes & Wasik, 1991; Fuson, 1990; Hiebert & Wearne, 1996). Simply using a correct procedure often does not lead to a better understanding of the underlying concepts. Further, at least one intervention study has cast doubt on the idea that procedural knowledge influences conceptual understanding. Children who received a lesson on the least common denominator procedure for fraction addition did not demonstrate increased conceptual understanding after the lesson (Byrnes & Wasik, 1991). This finding must be interpreted cautiously, both because the lesson was brief, and because children were near ceiling on the conceptual assessment at pretest. On the other hand, children and adults often have difficulty transferring instructed procedures to novel problems (see Singley & Anderson, 1989), suggesting that procedural knowledge is often not accompanied by conceptual understanding that allows people to adapt the procedure. Procedural knowledge may only lead to greater conceptual knowledge under certain circumstances, such as after extensive experience using the procedure, or when the relation between the procedure and the underlying concepts is relatively transparent.

As this review has shown, many studies provide suggestive evidence about the relations between conceptual and procedural knowledge. However, few studies have directly examined these relations. For several reasons, the methodologies commonly used in this area do not allow strong inferences to be drawn about how conceptual and procedural knowledge are related. First, in many studies, conceptual and procedural knowledge are not assessed independently. Consequently, it is impossible to ascertain how the two types of knowledge are related. Second, when independent assessments of conceptual knowledge are used, they are typically coarse-grained, categorical measures (i.e., children are classified as either having or not having the relevant conceptual understanding). Few studies have used fine-grained measures that can detect gradual changes in understanding. Third, in order to assess changes in knowledge, repeated assessments of children's knowledge are essential. Few studies in this area have assessed children's knowledge at more than one point in time. Finally, intervention studies are rare, and most do not include adequate controls that allow causal conclusions to be drawn.

The present study was designed to provide causal evidence about the relations between conceptual and procedural knowledge. The study used two measures of children's conceptual understanding—direct questions about the concepts and evaluation of unfamiliar procedures. Both types of measures have been used in past research to assess conceptual knowledge, so use of the two measures allowed us to compare alternative assessments and to obtain convergent validity for our findings. Both conceptual and procedural knowledge were assessed on two occasions—before and after instruction about either a concept or a procedure. Thus, the study should reveal whether instruction targeted at increasing one type of knowledge leads to gains in the other.

## Conceptual and Procedural Knowledge of Mathematical Equivalence

We addressed these issues in the context of children learning the principle of mathematical equivalence, which is the principle that the two sides of an equation represent the same quantity. Mathematical equivalence is a fundamental concept in both arithmetic and algebra. It incorporates at least three components: (a) the meaning of two quantities being equal, (b) the meaning of the equal sign as a relational symbol, and (c) the idea that there are two sides to an equation. These ideas are fundamental to mathematical problem solving; however, children's understanding of mathematical equivalence is often not challenged until children learn algebra. Most children in late elementary school do not seem to understand equivalence, as shown by their inability to solve novel problems such as $3 + 4 + 5 = 3 + \_$ (Alibali, 1999; Perry, Church, & Goldin-Meadow, 1988). Most third- through fifth-grade students either add all four numbers together or add the three numbers before the equal sign and ignore the final number.

This past research indicates that elementary school children have difficulty implementing correct procedures for solving mathematical equivalence problems. However, most previous research has not directly assessed children's understanding of the various components of the principle of equivalence. There is evidence that elementary school children interpret the equal sign as simply an operator signal that means "adds up to" or "produces" and do not interpret it as a relational symbol, meaning "the same as" (Baroody & Ginsburg, 1983; Behr, Erlwanger, & Nichols, 1980; Cobb, 1987; Kieran, 1981). Beyond this, it is unclear what aspects of equivalence children understand in late elementary school, what aspects they must learn in order to successfully solve mathematical equivalence problems, and what aspects they continue to misunderstand.

## Purpose of the Present Study

In summary, the goal of the present study was to provide causal evidence about the relations between children's conceptual and procedural knowledge of mathematical equivalence. More specifically, the study examined the impact of instruction about the concept of equivalence on children's problem-solving procedures and the impact of instruction about a problem-solving procedure on children's conceptual understanding of equivalence. In addressing these issues, we also identified what aspects of equivalence fourth and fifth graders understand, what aspects they learn easily, and what aspects they have difficulty learning.

To accomplish this goal, we assessed children's conceptual and procedural knowledge of mathematical equivalence both before and after instruction on either the concept of equivalence or on a correct procedure for solving equivalence problems, as well as in a no-instruction control group. We tested three specific hypotheses. First, we hypothesized that conceptual understanding and procedural skill are related. Thus, we predicted that children who generate correct procedures for solving the problems at the outset of the study will have greater conceptual understanding than children who generate incorrect procedures for solving the problems at the outset of the study. Second, we hypothesized that increasing children's conceptual knowledge will lead to gains in their procedural ability. Thus, we predicted that children who receive conceptual instruction will generate a correct procedure and transfer this procedure to related problems. Third, we hypothesized that increasing children's procedural knowledge will lead to gains in their conceptual understanding. Thus, we predicted that children who receive procedural instruction will improve their conceptual understanding. These predicted results would indicate that conceptual and procedural knowledge are yoked during the acquisition of mathematical equivalence.

## Method

### Participants

Participants were 60 fourth-grade students (24 girls and 36 boys) and 29 fifth-grade students (17 girls and 12 boys). The children were drawn from two suburban parochial schools that serve predominantly Caucasian, middle-income families. Three children were excluded from the sample for various reasons: an experimenter error was made while testing 1 student, 1 student was absent on the second day of testing, and 1 student had participated in a related study conducted at another school the previous year. The remaining sample of 86 children had a mean age of 10 years 2 months. Past research has shown that most fourth- and fifth-grade students are unable to solve equivalence problems correctly (Perry, 1991; Perry et al., 1988), so we did not expect performance to vary by grade level.

### Overview of Procedure

The study included a classroom screening, which took place before the experiment proper, and two experimental sessions. In the classroom screening, children were simply asked to solve two equivalence problems on a brief worksheet so we could identify children who did and did not solve the problems correctly. All children then participated individually in two experimental sessions. During the first session, all children completed two conceptual assessments. Then, children who solved the screening problems incorrectly and who were assigned to the instruction groups solved a set of pretest problems, received a brief lesson, and solved a set of posttest problems. On the following day, all children completed the conceptual assessments a second time, and then they completed a procedural assessment that included standard equivalence problems and transfer problems. Further details about the experimental procedure are provided below.

### Materials

### Conceptual Knowledge Assessments

*Question task.* We used a task analysis to identify three key components of the concept of equivalence: the meaning of two quantities being equal, the meaning of the equal sign, and the idea that an equation has two sides. Children answered seven questions designed to tap these three components. The items are described in Table 1.

*Evaluation task.* Children evaluated three correct and three incorrect procedures for solving mathematical equivalence problems (see Table 2 for examples). Evaluation of novel procedures is commonly used to measure children's conceptual knowledge (e.g., Gelman & Meck, 1983; Siegler & Crowley, 1994). Children were told that "students at another school solved these problems in lots of different ways," and they were then presented with examples of the procedures used by the students at the other school. Children were asked to evaluate each procedure as a "very smart, kind of smart, or not so smart" way to solve the problems. Each procedure was demonstrated twice in blocked random order. After children evaluated each procedure, they were asked to explain their reasoning for making that choice.

Table 1
*Items on the Question Task*

| Concept | Item | Correct response |
|---|---|---|
| Meaning of equal quantities | Define what it means for two sets of objects to be equal | Mention "the same" or "the same number" |
| | Identify a pair of numbers that is equal to a given pair (e.g., 5 + 3) | Choose correct pair |
| Meaning of the equal sign | Define the equal sign (with the prompt, "can the equal sign mean anything else?") | Mention "the same" or "equal" in the definition |
| | Rate 2 complete and 4 incomplete or incorrect definitions of the equal sign as "very smart, kind of smart, or not so smart" definitions | Rate both complete definitions as "very smart" |
| Structure of equation | Reproduce 2 standard addition problems (e.g., 4 + 3 + 6 + 4 = __) and 2 equivalence problems (e.g., 5 + 4 + 7 = 5 + __) from memory after a 5 s delay (to assess encoding) | Correctly reproduce the position of the equal and plus signs in all 4 problems |
| | Identify the two sides of a mathematical equivalence problem | Correctly identify the addends before the equal sign as one side, and the addends and blank after the equal sign as the other side |
| Meaning of the equal sign and structure of equation | Indicate whether 15 problems made sense or not—2 legitimate and 1 illegitimate of each of the following forms: standard (3 + 4 = 7), reverse (8 = 2 + 6), symmetrical (2 + 5 = 5 + 2), alternative (3 + 2 = 6 − 1), or identity (5 = 5; to assess recognition of the use of the equal sign in multiple contexts; adapted from Baroody & Ginsburg, 1983; Cobb, 1987) | Appropriately evaluate more than 9 out of the 12 nonstandard problems as making sense or not |

Table 2
*Commonly Used Procedures for Solving Equivalence Problems*

| Procedure | Explanation used by experimenter in the evaluation task | Sample explanation given by a child during the pretest or posttest |
|---|---|---|
| Correct procedures | | |
| Equalize | She added the 3, the 4, and the 5 together and got an answer, and then she figured out what number she needed to add to this 3 (point) to get that same answer. | "3 plus 4 plus 5 is 12, and 3 plus 9 is 12." |
| Add subtract | She added the 3, the 4, and the 5 together, and then subtracted this 3 (point). | "I added 3 and 4 and 5, and then took away 3." |
| Grouping | Because there is a 3 here (point) and a 3 here (point), she only added the 4 and the 5 together. | "I added 4 plus 5." |
| Incorrect procedures | | |
| Add all | She added the 3, the 4, the 5, and the 3 together. | "3 plus 4 is 7, plus 5 is 12, plus 5 more is 17." |
| Add to equal sign | She added the 3, the 4, and the 5 together. | "I added 3 plus 4 plus 5 and it made 12." |
| Carry | Because there was a 5 here, she wrote a 5 in the blank. | "I just wrote the 5 in the blank." |

*Note.* All explanations are for the problem $3 + 4 + 5 = 3 + \_\_$. All of the explanations used by the experimenter in the evaluation task included gestures that conveyed the same information as the accompanying speech. In the examples above, gestures are noted only where needed to disambiguate the referents of speech.

## *Procedural Knowledge Assessments*

All problems were printed on legal size paper and laminated so that children could write on the problems with markers, and their answers could later be erased. Problems were presented on a table-top easel.

*Standard equivalence problems.* All screening, pretest, training, and posttest problems were of the form $a + b + c = a + \_\_$.

*Transfer problems.* Transfer problems were constructed by varying three features of the problems: the operation used (addition or multiplication), the position of the blank (in the final position or immediately following the equal sign), and the presence of equivalent addends on both sides of the equation (yes or no). This yielded five new types of problems, as well as the standard equivalence problem type. The transfer test included two instances of each type of problem, for a total of 12 problems. The breakdown of problem types is presented in Table 3.

# *Procedure*

## *Classroom screening*

A few days before data collection began, classroom teachers administered a brief paper-and-pencil assessment that consisted of two standard equivalence problems. Teachers did not mention that these problems were related to this study. This classroom screening was used to identify children who solved the problems correctly (equivalent children; $n = 27$) and children who solved the problems incorrectly (nonequivalent children; $n = 59$). Nonequivalent children were then randomly assigned to one of three groups: conceptual instruction, procedural instruction, or control (no instruction). All children then participated individually in two sessions that lasted approximately 15–25 min each. Sessions were held on 2 consecutive days.

## *Session 1*

Children completed the two conceptual assessments (described above) in counterbalanced order.[1] Equivalent children and nonequivalent children who had been assigned to the control group ended the first session after completing the conceptual assessments.

Children in the instruction groups went on to solve four standard equivalence problems in a pretest. The problems were presented on the table-top easel, and the children were asked to solve each problem however they thought was best. When the children finished solving each problem, the experimenter asked, "How did you get that answer?" and the children explained how they solved the problem. On the final pretest problem, the experimenter told the children whether they had solved the problem correctly. Unexpectedly, 11 children solved the final pretest problem (and in some cases, some of the other pretest problems) correctly, so they were not provided with instruction.[2] This left 17 children in the conceptual-instruction group, 15 children in the procedural-instruction group, and 16 children in the control group.

The children in the instruction groups then received a brief lesson. In the conceptual-instruction group, a problem was presented, and the children were told the principle behind the problems. The instruction was based on that used by Perry (1991) and Alibali (1999) and included both speech and appropriate gestures. The children were told the following:

> Because there is an equal sign (point to the equal sign), the amount before it (sweep hand under left side of equation) needs to equal the amount after it (sweep hand under right side of equation). That means that the numbers after the equal sign (sweep hand under right side of equation) need to add up to the same amount as the numbers before the equal sign (sweep hand under the left side of the equation).

No instruction was given in any procedure for solving the problem,

---

[1] The encoding item from the question task was always administered first, regardless of the order of the two conceptual assessments, to ensure that experience with problems from the evaluation task did not bias children's performance on the encoding item.

[2] These children had slightly higher scores on the question and evaluation tasks than the other children who were categorized as nonequivalent on the basis of the classroom screening. They did not differ systematically from the nonequivalent children in gender distribution or grade level. We believe that this group included both children who were careless on the classroom screening problems (and thus were really equivalent) and children who learned from the conceptual assessments. Because the group was small and heterogeneous, we did not analyze them as a separate group.

Table 3
*Problems on the Transfer Test*

| Equivalent operands | Operation | Position of blank | Example form |
|---|---|---|---|
| Yes | Addition | Right | $a + b + c = a + \_\_$ (standard) |
| Yes | Addition | Left | $a + b + c = \_\_ + c$ |
| Yes | Multiplication | Right | $a \times b \times c = a \times \_\_$ |
| Yes | Multiplication | Left | $a \times b \times c = \_\_ \times c$ |
| No | Addition | Right | $a + b + c = d + \_\_$ |
| No | Addition | Left | $a + b + c = \_\_ + d$ |

and no solutions were provided.[3] In the procedural-instruction group, the problem was presented, and children were taught the grouping procedure. They were told the following:

> There is more than one way to solve these problems, but one way is like this. Because there is a (number) here and a (number) here (point to the repeated addends), all you need to do is add the (number) and the (number) together (point to the nonrepeated addends).

After instruction, children were asked to solve a new problem. After solving this problem (without feedback), the instruction was repeated and the children solved another problem. If the children requested help, the appropriate instructional statements were repeated. After the instruction phase, children solved and explained four posttest problems identical in form to the pretest and training problems.

## *Session 2*

Session 2 was identical for all of the children. First, children completed the two conceptual assessments in the same counterbalanced order as on the previous day. Then children solved and explained the 12 problems on the transfer test, which were presented in one of two random orders. Although equivalent children participated in this session, their results from Session 2 are not reported in this article because they are not germane to the issues addressed here.

At the end of the experimental session, all children received a brief lesson on the concept of equivalence and on a correct procedure for solving the problems, to ensure that every child who participated in the study benefited from their participation. Before returning to class, the children were allowed to choose a brightly colored pencil as a reward for their participation.

## *Coding*

*Question task.* Each of the seven items was scored as a success (1 point) or not, for a possible question score of 7 (see Table 1).[4]

*Evaluation task.* For each demonstrated procedure, ratings of "very smart" were scored as 2, ratings of "kind of smart" were scored as 1, and ratings of "not so smart" were scored as 0. To calculate each child's evaluation score, we calculated each child's mean rating for the three correct procedures and for the three incorrect procedures, and we took the difference between these two means. This score is a measure of how much each child differentiated between correct and incorrect procedures. Evaluation scores could range from $-2.0$ (indicating that children rated incorrect procedures higher than correct procedures) to $+2.0$ (indicating that children rated correct procedures higher than incorrect procedures).

*Problem-solving tasks.* The procedures children used to solve the pretest, posttest, and transfer problems were identified (see Table 2 for examples). In many cases, children's solutions to the problems revealed the procedures they had used to solve the problems (e.g., for the problem $3 + 4 + 5 = 3 + \_\_$, the solution 15 reveals that the child used the add-all strategy). When a solution was ambiguous, the child's explanation of the procedure was used to identify the problem-solving procedure. All results are based on correct or incorrect procedures, rather than solutions, because children sometimes made arithmetic errors.

## *Reliability*

Reliability was assessed for 19 randomly selected participants in Session 1 and 19 different participants in Session 2. An independent rater coded the entire session from videotape. The correlation between raters on question scores was .92, and agreement on individual items ranged from 87% to 100%. The correlation between raters on evaluation scores was .99. Agreement between raters was 99% for assigning procedure codes on the pre- and posttests, 97% for assigning procedure codes on the transfer test, and 100% for assigning accuracy scores to the pretest, posttest, and transfer test.

# Results

The results are organized around three key issues: (a) the relation between conceptual and procedural knowledge at the outset of the study, (b) changes in conceptual knowledge due to instruction, and (c) changes in procedural knowledge due to instruction. Gender and grade effects were rare and are noted where applicable. Unless otherwise noted, differences were significant at $p < .05$.

## *Relation Between Initial Conceptual and Procedural Knowledge*

We first examined the relation between conceptual and procedural knowledge at the outset of the study. As ex-

[3] There are many types of possible conceptual instruction, and we chose to teach the concept behind the specific problems rather than an abstract concept. Some might argue that a procedure was implicit in our instruction; however, the instruction did not specify how to carry out any procedure. Although there may be commonalities between our two instructional conditions, important differences exist as well.

[4] Of the children who recognized that problems such as $7 + 2 = 2 + 7$ made sense, 62% ($n = 43$) also thought that the problem $5 - 3 = 3 - 5$ made sense. This suggests that many children overextend the addition principle of commutativity to subtraction.

pected, children who solved at least one of the classroom screening problems correctly (equivalent children) had greater conceptual understanding in Session 1 than children who solved both of the classroom screening problems incorrectly (nonequivalent children). Equivalent children had higher scores than nonequivalent children on both the question task (*M*s = 3.74 vs. 2.95, out of 7), $t(84) = 2.68$, and the evaluation task (*M*s = 1.13 vs. 0.32, possible range −2 to +2), $t(84) = 6.50$. Success on the question and evaluation tasks was not influenced by the order in which the assessments were administered. Neither grade nor gender was related to success on the classroom screening problems or on the evaluation task. However, on the question task, fifth graders scored higher than fourth graders (*M*s = 3.7 vs. 2.9), $t(84) = 2.61$, and girls scored higher than boys (*M*s = 3.5 vs. 2.9), $t(84) = 2.09$.

Success on the question and evaluation tasks was moderately positively correlated ($r = .37$). Children were classified as having scores above or below the median on each measure, and the association between the measures was assessed. Children who scored highly on one measure tended to score highly on the other measure, $\chi^2(1, N = 86) = 9.81$. Children were further classified as scoring above the median on both, one, or neither measure. Equivalent children tended to have high scores on both measures, whereas nonequivalent children tended to have low scores on both measures, $\chi^2(2, N = 86) = 14.58$.

Table 4 presents the percentage of equivalent and nonequivalent children who succeeded on each of the seven items on the question task at the outset of the study. As seen in the table, most children in both groups understood what it meant for two quantities to be equal. Compared to nonequivalent children, equivalent children were more likely to succeed on items designed to tap children's understanding of the meaning of the equal sign and of the structure of equations. Nevertheless, many children solved equivalence problems correctly even though they did not understand all three key components of equivalence.

Table 5 presents the mean ratings for individual procedures on the evaluation task for equivalent and nonequivalent children. Equivalent children rated all of the correct procedures more positively than they rated the incorrect procedures. This was not due to different children rating only one correct procedure highly; 74% of equivalent children rated more than one correct procedure as "very smart." Nonequivalent children, on the other hand, rated both correct and incorrect procedures highly. They rated the grouping procedure less positively, and the add-all and add-to-equal-sign procedures more positively, than equivalent children. However, like the equivalent children, nonequivalent children rated the equalize procedure highest overall. This suggests that they recognized a correct procedure before they used one. However, this interpretation must be made cautiously because children's explanations of their ratings sometimes suggested that they misinterpreted the procedure (e.g., children occasionally said that the equalize procedure was very smart because the child had added all of the numbers).

Overall, these results indicate that children who had greater procedural knowledge, as shown by their ability to solve the problems correctly on the classroom screening, also had greater conceptual knowledge of equivalence. However, conceptual understanding was not all-or-none; children who solved the problems correctly often did not have a complete understanding of equivalence. Further, the question and evaluation tasks seemed to tap related, but not identical, aspects of understanding of equivalence. Finally, these results support the validity of the conceptual measures. Both measures distinguished between children who were and were not able to solve the problems correctly on their own.

### *Changes in Conceptual Knowledge*

Nonequivalent children were expected to improve on the conceptual assessments after receiving instruction. Figures 1 and 2 display the average score for each group in each session for the question task (see Figure 1) and the evaluation task (see Figure 2). At Session 1, there were no group differences on either measure of conceptual understanding.

Table 4
*Percentage of Children in Each Knowledge Group Who Succeeded on Each Item on the Question Task in Session 1*

| Concept and task | Group: Equivalent | Group: Nonequivalent |
|---|---|---|
| Meaning of equal quantities | | |
| Define equal | 93 | 92 |
| Identify equal pairs of numbers | 96 | 92 |
| Meaning of the equal sign | | |
| Define the equal sign | 30 | 19 |
| Rate correct definitions of equal sign as "very smart" | 15 | 17 |
| Structure of equation | | |
| Encode equations correctly | 56 | 36† |
| Identify the 2 sides of an equation | 33 | 22 |
| Meaning of the equal sign and structure of equation | | |
| Recognize use of equal sign in multiple contexts | 52 | 19* |

†$p < .10$, *$p < .01$, chi-square tests of the difference between equivalent and nonequivalent children.

Table 5
*Mean Rating of Each Procedure in Session 1 by Knowledge Group*

| Procedure | Group | |
|---|---|---|
| | Equivalent | Nonequivalent |
| Correct procedures | | |
| Equalize | 1.57 | 1.47 |
| Add subtract | 1.15 | 1.00 |
| Grouping | 1.39 | 0.70* |
| Incorrect procedures | | |
| Add all | 0.28 | 1.26* |
| Add to equal sign | 0.37 | 0.80* |
| Carry | 0.07 | 0.14 |

*Note.* Very smart = 2, kind of smart = 1, not so smart = 0.
$^*p < .01$, *t* tests of difference between equivalent and nonequivalent children.

To evaluate learning, we calculated the change on each measure from Session 1 to Session 2 for each child. Then, we conducted two orthogonal planned comparisons with each measure. We first tested for an effect of instruction (regardless of type), and then tested for an effect of instruction type (conceptual vs. procedural).

### *Question Task*

Children who received instruction improved more on the question task than children who did not receive instruction, $F(1, 45) = 5.57$ (see Figure 1). Children in the conceptual-instruction group improved more on the question task than children in the procedural-instruction group (mean increase = 1.18 vs. 0.67); however, the amount of change did not differ significantly across the two instruction groups, $F(1, 45) = 2.40$, $p = .13$. The same pattern is observed if one assesses the number of children in each group who improved on the question task: 76% of children in the conceptual-instruction group, 53% of children in the procedural-instruction group, and 38% of children in the control group had higher question scores in Session 2 than in Session 1. At this individual level, more children in the instruction groups than in the control group tended to improve on the question task, $\chi^2(1, N = 48) = 3.43, p = .06$, but type of instruction was not significantly related to improvement, $\chi^2(1, N = 32) = 1.73$. Thus, both conceptual and procedural instruction led to increased conceptual understanding, as assessed with the question task.

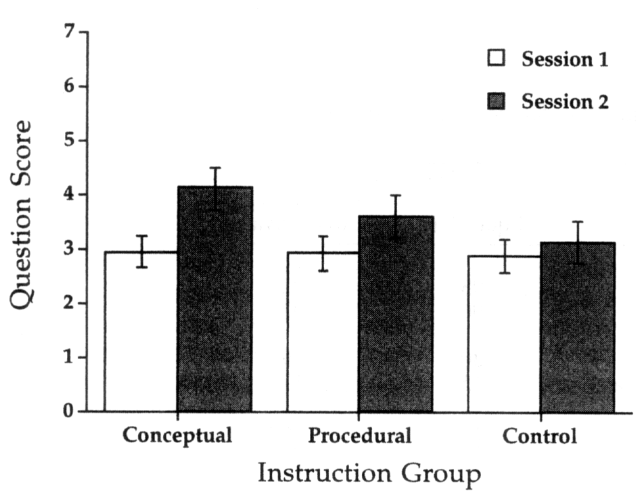


*Figure 1.* Average question scores (with error bars representing standard errors) for each instruction group at each session.

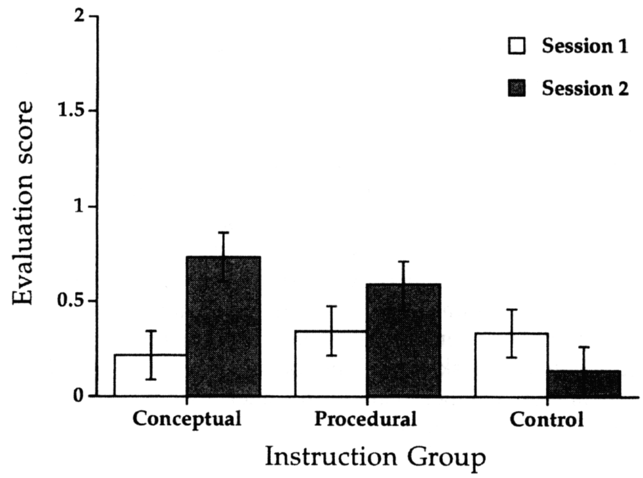


*Figure 2.* Average evaluation scores (with error bars representing standard errors) for each instruction group at each session.

As shown in Table 6, both types of instruction seemed to help children recognize the full meaning of the equal sign and the use of the equal sign in multiple contexts. Conceptual instruction was somewhat more effective than procedural instruction at helping children to understand the structure of equations. Overall, instruction did not lead to uniform change on a single item; different children learned different aspects of equivalence from the same instruction.

### *Evaluation Task*

Parallel results were found for the evaluation task (see Figure 2). Children in the instruction groups improved on the evaluation task, whereas children in the control group worsened on the task, $F(1, 45) = 7.57$. Again, there was a trend for children in the conceptual-instruction group to change more than children in the procedural-instruction group (mean change = +0.52 vs. +0.23); however, this difference did not reach significance, $F(1, 45) = 1.37$. Again, the same pattern is observed if one assesses the number of children in each group who improved on the task: 71% of children who received conceptual instruction, 53% of children who received procedural instruction, and 31% of children who received no instruction improved on the evaluation task. At this individual level, more children in the instruction groups than in the control group improved on the evaluation task, $\chi^2(1, N = 48) = 4.17$, but type of instruction was not significantly related to improvement, $\chi^2(1, N = 32) = 0.95$. Thus, both conceptual and procedural instruction led to increased conceptual understanding, as assessed with the evaluation task.

Table 6
*Proportion of Children in Each Instruction Group Who Succeeded on Items on the Question Task in Session 2, After Failing on the Item in Session 1*

| Concept and task | Conceptual instruction | Procedural instruction | Control |
|---|---|---|---|
| Meaning of the equal sign | | | |
| Define the equal sign | .33 (12) | .25 (12) | .23 (13) |
| Rate correct definitions of equal sign as "very smart" | .23 (13) | .18 (11) | .00 (15) |
| Structure of equation | | | |
| Encode equations correctly | .58 (12) | .30 (10) | .30 (10) |
| Identify the 2 sides of an equation | .14 (14) | .00 (11) | .00 (12) |
| Meaning of the equal sign and structure of equation | | | |
| Recognize use of equal sign in multiple contexts | .33 (15) | .29 (14) | .08 (13) |

*Note.* Items assessing the meaning of equal quantities are not presented because children were at ceiling on these items in Session 1. The numbers of children who failed each item in Session 1 are in parentheses.

Table 7 presents the mean ratings for individual procedures for children in each group. Children in the control group decreased their ratings of correct procedures while maintaining their ratings of incorrect procedures. In contrast, children in the conceptual-instruction group increased their ratings of two correct procedures and decreased their ratings of two incorrect procedures. Children in the procedural-instruction group greatly increased their rating of the grouping procedure, while slightly decreasing their ratings of most of the other procedures, including other correct procedures.

These data suggest that the increases in evaluation scores for the procedural-instruction group were due almost exclusively to recognizing the instructed procedure, grouping, as correct in Session 2. To explore this possibility, we removed each child's rating of the grouping procedure and recalculated each evaluation score. Although evaluation scores still increased for the conceptual-instruction group (mean change = +0.54), evaluation scores decreased for both the procedural-instruction and control groups (mean change = −0.26 and −0.17, respectively). There was no longer an overall effect for receiving instruction, $F(1, 45) = 1.01$, and children in the conceptual-instruction group outperformed children in the procedural-instruction group, $F(1, 45) = 11.94$. These results suggest that improvements in evaluation scores for the procedural-instruction group were primarily due to children rating the instructed procedure higher in Session 2 (while decreasing their rating of other correct procedures), and may not reflect true increases in conceptual understanding.

## *Changes in Procedural Knowledge*

### *Learning New Procedures From Instruction*

Most of the children in both instruction groups used a correct procedure on the posttest (see Table 8). In the procedural-instruction group, every child but one adopted the instructed procedure on the posttest, and these children used the instructed procedure on all four posttest problems. In the conceptual-instruction group, more than half of the children generated the equalize procedure on at least one posttest problem, about a third of the children generated the grouping procedure, and one child generated the add-subtract procedure. This variability in procedure generation by children in the conceptual-instruction group supports our claim that the conceptual instruction did not directly teach a specific procedure. Overall, children in the conceptual-instruction group used a larger variety of procedures on the posttest than did children in the procedural-instruction group ($Ms$ = 1.29 vs. 1.00), $t(30) = 2.42$, but children in the two groups solved an equivalent number of problems correctly on the posttest (and on the two standard problems presented in Session 2).

Table 7
*Mean Change in Evaluation of Each Procedure for Each Instruction Group*

| Procedure | Conceptual instruction | Procedural instruction | Control |
|---|---|---|---|
| Correct procedures | | | |
| Equalize | +0.53 (1.76) | −0.43 (1.23) | −0.09 (1.22) |
| Add subtract | −0.09 (0.94) | −0.47 (0.40) | −0.22 (0.78) |
| Grouping | +0.15 (0.82) | +1.03 (1.87) | −0.28 (0.34) |
| Incorrect procedures | | | |
| Add all | −0.79 (0.53) | −0.30 (0.97) | +0.06 (1.28) |
| Add to equal sign | 0.00 (0.76) | −0.30 (0.70) | −0.03 (0.63) |
| Carry | −0.18 (0.06) | +0.03 (0.10) | 0.00 (0.06) |

*Note.* The numbers in parentheses are the mean evaluations in Session 2.

Table 8
*Percentage of Children in the Conceptual- and Procedural-Instruction Groups Who Used Correct Procedures on the Posttest*

| Instruction group | Procedure | | | |
|---|---|---|---|---|
| | Grouping | Equalize | Add subtract | Any correct |
| Conceptual | 29 | 59 | 6 | 82 |
| Procedural | 93 | 0 | 0 | 93 |

*Note.* Children could use more than one correct procedure on the posttest.

### *Success on Transfer Problems*

Children in the conceptual-instruction group succeeded on the greatest number of transfer problems ($M$ = 5.9 out of 10), followed by children in the control group ($M$ = 2.7) and children in the procedural-instruction group ($M$ = 1.7). As these scores suggest, there was not an overall effect for instruction, but there was an effect for instruction type, $F(1, 45) = 9.30$. Order of presentation of the transfer problems did not influence success. The surprising finding that children in the control group outperformed children in the procedural-instruction group was due to the fact that four children in the control group solved all 10 of the transfer problems correctly, whereas none of the other children in the control group solved more than two transfer problems correctly ($M$ = 0.25).[5] With these 4 children omitted from the analysis, there was both a positive effect of instruction, $F(1, 45) = 11.90$ and an effect for instruction type, $F(1, 45) = 14.84$.

To explore transfer performance further, we examined differences on specific problem types. Because we were primarily interested in differences due to type of instruction, and because of the ambiguities in the control group, we focused on differences between the two instruction groups. As shown in Figure 3, the instruction groups differed on four out of the five types of transfer problems. On problems that only required adapting the procedure to a new operation (multiplication), the two groups were equally successful. On all of the other transfer problem types, children in the conceptual-instruction group solved more problems correctly than children in the procedural-instruction group.

### *Procedure Use on the Transfer Problems*

Children who received conceptual instruction tended to use correct procedures on transfer problems (most often equalize), whereas children who received procedural instruction usually reverted to their old incorrect procedures or switched to other common incorrect procedures (e.g., change from using add all to using add to equal sign; see Figure 4). Only 13% of children in the procedural-instruction group attempted to apply a correct procedure to all of the transfer problems, compared to 59% of children in the conceptual-instruction group, $\chi^2(1, N = 32) = 7.04$. Further, 29% of children in the conceptual-instruction group used more than one correct procedure on the transfer test, whereas none of the children in the procedural-instruction group did, $\chi^2(1, N = 32) = 5.23$. As seen in Figure 4, the poor transfer in the procedural-instruction group did not result from faulty adaptation of a correct procedure, such as simply adding the second and third addends even when the other two addends were not the same (e.g., adding 4 + 3 in problem 7 + 4 + 3 = __ + 3); such attempts occurred on fewer than 5% of trials. Group differences also were not simply attributable to differences in which correct procedure children used. Children in the procedural-instruction group adopted the grouping procedure ($N$ = 14) and solved an average of fewer than two transfer problems correctly, whereas children in the conceptual-instruction group who generated only the grouping procedure on the posttest ($N$ = 4) solved an average of five transfer problems correctly, $t(16) = 1.54, p = .14$.

In summary, children in both instruction groups learned correct procedures. However, learning in the procedural-instruction group was overly specified and narrow. Rather than attempt to adapt their new, correct procedure to novel problems, children in the procedural-instruction group simply reverted to incorrect procedures. In contrast, children in the conceptual-instruction group generalized their new knowledge much more broadly.

## Discussion

The present study examined the relations between conceptual and procedural knowledge of mathematical equivalence. By assessing children's knowledge both before and after randomly assigned instruction, we obtained causal evidence that conceptual and procedural knowledge influence one another. The findings from this study hinged on the use of adequate and repeated assessments of both conceptual and procedural knowledge. Instead of using categorical knowledge assessments, which cannot detect gradual changes in children's understanding, we used multiple, continuous measures of both conceptual and procedural knowledge. Using these measures, we found converging results about the interconnections between the two types of knowledge. We first discuss these relations and then explore potential mechanisms that may underlie them. Finally, we consider implications of this work for education.

### *Relations Between Conceptual and Procedural Knowledge*

The present results highlight the causal, bidirectional relations between conceptual and procedural knowledge in children learning mathematical equivalence. Children who received conceptual instruction not only increased their

[5] Children in the control group did not solve the pretest problems presented during the experimental session, so these 4 children may be similar to the children who were originally assigned to an instruction group but who then solved the final pretest problem correctly (see Footnote 2). Indeed, the proportion of children assigned to the instruction groups who solved the final pretest problem (and most of the transfer problems) correctly was comparable to the proportion of children in the control group who succeeded on the transfer problems.

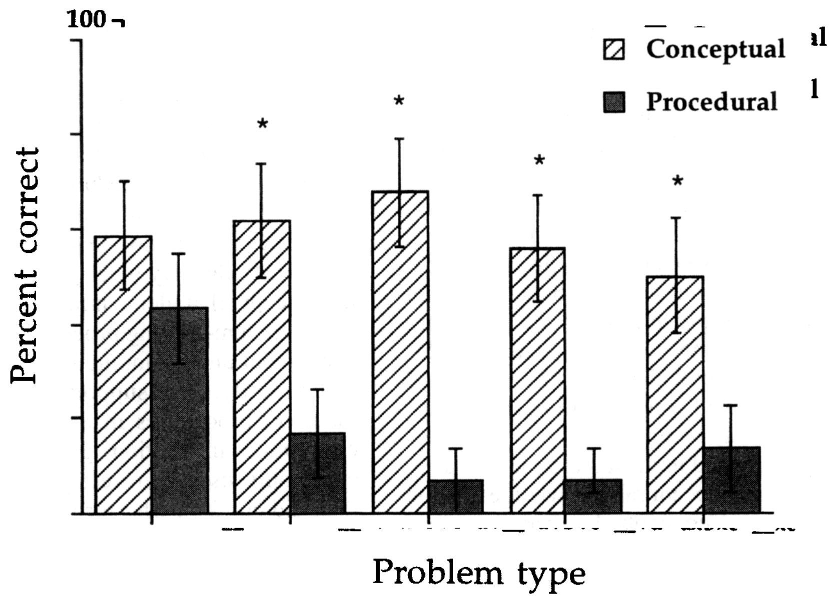


*Figure 3.* Percentage correct (with error bars representing standard errors) on each type of transfer problem for children in the conceptual- and procedural-instruction groups. Asterisks indicate significant differences between the groups.

conceptual understanding, but also generated several correct, flexible problem-solving procedures. Children who received procedural instruction not only adopted a correct problem-solving procedure, but also increased their conceptual understanding. Thus, conceptual understanding can lead to procedure generation, and procedural knowledge can lead to gains in conceptual understanding.

These results suggest that there is an iterative relationship between conceptual and procedural knowledge. Increases in one type of knowledge can lead to gains in the other type of knowledge, which in turn may lead to further increases in the first. Thus, at least under some circumstances, conceptual and procedural knowledge develop in tandem, rather than independently.

However, although conceptual and procedural knowledge influence one another, the strength of their influence may not be symmetrical. Children who received procedural instruction showed greater gains on both the question and evalua-

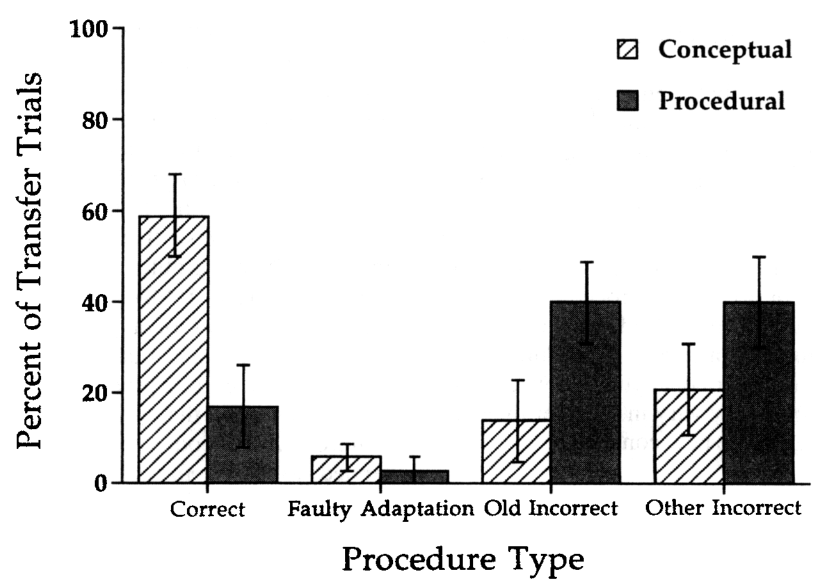


*Figure 4.* Percentage of transfer trials (with error bars representing standard errors) on which each type of procedure was used by children in the conceptual- and procedural-instruction groups.

tion tasks than children in the control group, but these gains were modest and were smaller than the gains made by children who received conceptual instruction. Children in the procedural-instruction group also had significantly poorer transfer performance than children in the conceptual-instruction group. They often did not try to use the instructed procedure on problems that differed only in small ways from the instructional problems, a finding that is consistent with previous research on transfer performance (Singley & Anderson, 1989). The children did not seem to have sufficient conceptual understanding to apply and adjust the procedure to novel problems. Thus, outcomes on the question, evaluation, and transfer tasks all suggest that gains in procedural knowledge *can* lead to improved conceptual understanding, but that such improvement may be limited.

In contrast, gains in conceptual understanding led to fairly consistent improvements in procedural knowledge in this study. Children who received conceptual instruction were just as likely to learn a correct procedure as children who received procedural instruction, and conceptual instruction led to better transfer performance than procedural instruction. Thus, conceptual knowledge seemed to have a greater impact on procedural knowledge than the reverse.

Past research supports the idea that there is an asymmetric relationship between conceptual and procedural knowledge. Perry (1991) also found that conceptual instruction was more effective than procedural instruction at promoting broad transfer. Further, studies have shown that most children who understand fundamental concepts within a domain are able to use a correct procedure for solving related problems (e.g., Briars & Siegler, 1984; Cauley, 1988; Cowan & Renton, 1996), whereas children who learn procedures sometimes never master the concepts behind them (e.g., Fuson, 1990; Kouba, Carpenter, & Swafford, 1989). For example, Hiebert and Wearne (1996) found that conceptual understanding predicted future procedural skill, but that a significant number of children who adopted a correct procedure did not learn the related concepts over the course of 3 years.

An asymmetric relationship between the two types of knowledge may be due in part to individual differences in the use of procedural knowledge for generating conceptual knowledge. In our study, about half of the children who received procedural instruction also improved in their conceptual understanding of the domain. This suggests that some children try to figure out why procedures work, whereas others are content to use a procedure without understanding it. In comparison, most children who received conceptual instruction generated a correct procedure. Thus, children may commonly use their conceptual knowledge to generate procedures, but they may vary in whether they use procedural knowledge to generate new concepts.

### *Mechanisms Through Which Conceptual and Procedural Knowledge Influence One Another*

The present findings demonstrate that conceptual and procedural knowledge influence one another. However, they leave open the question of *how* each type of knowledge influences the other. The data provide some clues about possible mechanisms.

First, consider how conceptual knowledge might influence procedure generation. One possibility is that gains in conceptual knowledge help children to recognize that their problem-solving procedures are incorrect. As children solidify their conceptual knowledge, they may begin to realize that their problem-solving procedures are inconsistent with that knowledge. By highlighting such inconsistencies, gains in conceptual understanding may lead children to change their problem-solving procedures. In this study, the idea that conceptual understanding helps children to recognize incorrect procedures is supported by children's evaluations of incorrect procedures in Session 2. Children who received conceptual instruction gave lower ratings to the most commonly used incorrect procedure, add all, in Session 2 than did children who received procedural instruction or no instruction (see Table 7). Thus, conceptual knowledge may influence procedure generation by helping children recognize the need for new procedures.

A second possibility is that conceptual knowledge constrains procedure generation. That is, conceptual knowledge may prevent children from generating incorrect procedures. In this study, most children in the conceptual-instruction group generated correct procedures—ones that were consistent with the concepts presented in the instruction. Indeed, the most commonly generated procedure, equalize, had the most straightforward mapping to the conceptual instruction. Some children also generated additional correct procedures when faced with novel transfer problems. Children in the conceptual-instruction group also recognized multiple correct procedures and distinguished them from incorrect procedures on the evaluation task in Session 2 (see Table 7). Thus, conceptual knowledge may influence procedure generation by helping children to identify essential elements of correct procedures and by helping them to monitor whether potential procedures are worth trying.

Both of these explanations for the influence of conceptual knowledge on procedure generation could be driven by another underlying mechanism: Conceptual knowledge could lead to changes in children's encoding of the problems. Correct encoding of problems is an essential step in successful problem solving (Alibali, 1998; Siegler, 1976), and recent work suggests that changes in encoding lead children to generate new problem-solving procedures (Alibali, McNeil, & Perrott, 1998). Indeed, correctly identifying the important features of a problem may be the first step in recognizing the inadequacy of old procedures and in constructing new, correct procedures. In the current study, many children encoded the problems incorrectly at Session 1, and children who encoded the problems correctly at Session 1 tended to solve the problems correctly (see Table 4). Furthermore, children in the conceptual-instruction group demonstrated more change on the encoding item of the question task than on any other item, and improvement on this item differentiated them the most from the procedural-instruction and control groups (see Table 6). Thus, conceptual knowledge may influence procedure generation by highlighting features of the problems that children need to encode.

Next, consider how procedural knowledge might influence conceptual understanding. One possibility is that procedural knowledge constrains conceptual understanding. Knowledge of correct procedures may help children to zero in on accurate conceptual knowledge. For example, correct procedures for solving math equivalence problems are inconsistent with naive conceptions of the equal sign as meaning "get the answer" and as signaling the end of the problem. In this study, procedural instruction led to improved understanding of the equal sign and its implications for the structure of equations (see Table 6). Thus, procedural knowledge could influence conceptual understanding by helping to eliminate misconceptions.

A second possibility is that, when children use procedures, they sometimes think about why they work. When procedures are easy to implement (such as the grouping procedure in this study), children may not use all their resources in implementing the procedure, and they may have resources available to consider the basis of the procedure. Further, in situations in which newly learned procedures result in different solutions than prior procedures, children may find this surprising, and this may lead them to consider the conceptual basis of the new procedure.

There may also be individual differences in children's tendency to consider the conceptual basis of the procedures they use. In this study, about half of the children who received procedural instruction improved on each measure of conceptual understanding. A similar result was found for children learning multidigit arithmetic (Hiebert & Wearne, 1996). These individual differences are reminiscent of individual differences among college students in their tendencies to explain things spontaneously to themselves as they read example problems (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Pirolli & Recker, 1994; Renkl, 1997). In these studies, students who self-explained were more successful on posttest problems than those who did not self-explain. Students who self-explained were also more likely to recognize when they did not understand something, and this recognition may have motivated them to try to explain the reasoning behind the problems. Similar individual differences in self-explanation may help to account for variability in how much children learn from procedural instruction. It is possible that procedural knowledge influences conceptual understanding only for those children who attempt to self-explain why particular procedures are correct. If this is the case, then prompting children to explain why procedures are correct may help to increase the impact of procedural knowledge on conceptual knowledge. In support of this view, recent research indicates that prompting students to generate explanations while they study leads to improved learning (Chi, de Leeuw, Chiu, & LaVancher, 1994; Siegler, 1995).

### *Educational Implications*

The present findings suggest that by fourth grade, most children understand what it means for two quantities to be equal. However, most children in late elementary school do not fully understand the meaning of the equal sign or the structure of equations. It is likely that children's simplified conception of the meaning and role of the equal sign develops because they see thousands of problems with an equal sign just before the answer and few with the equal sign in other positions or contexts (Baroody & Ginsburg, 1983). Even after exposure to the use of the equal sign in multiple contexts, many children still do not extract the full, relational meaning of the equal sign. This simplified understanding may become a serious handicap when children are introduced to algebra (see Kieran, 1981; Sfard & Linchevski, 1994). Indeed, without a prior understanding of equivalence, algebraic equation-solving procedures may not make sense. Children often have difficulties solving symbolic algebra problems (Koedinger & MacLaren, 1997), and their incomplete understanding of the meaning and role of the equal sign may be one source of these difficulties.

These findings also have broader implications for the content of mathematics lessons. In this study, conceptual instruction led to the greatest gains in conceptual understanding and to the most transferable problem-solving skills. Children who received conceptual instruction sometimes generated multiple procedures, and they were able to adapt their procedures to novel problems. Procedural instruction led to more modest gains in both understanding and problem-solving ability. When children were directly taught a procedure, most children used the procedure on problems that were formally identical to those used in instruction, but they did not attempt to apply it to problems with minor variations in surface features. Similar transfer results were observed by Perry (1991) for children who were taught a different procedure for solving mathematical equivalence problems (add subtract), suggesting that these results are not limited to a particular procedure. Taken together, these findings suggest that, under certain circumstances, children may benefit most from conceptual instruction that helps them to invent correct procedures on their own. When children are empowered with fundamental concepts, they are able to solve novel problems on their own. Our results are in line with current reforms in mathematics education, which recommend emphasizing the conceptual underpinnings of tasks (NCTM, 1989).

However, at least five features of the current study should be considered before generalizing these findings beyond the instructions, task, and sample used in this study. First, the conceptual instruction given in this study was closely tied to the target problems. More abstract instruction may not be sufficient for procedure generation. Second, correct procedures for solving equivalence problems are not overly complex or difficult for children of this age (note that some children generated a correct procedure before the intervention). It is possible that conceptual instruction may facilitate procedure generation only under these conditions. Third, instruction on procedures with more transparent mappings to the underlying concepts or instruction on a procedure, along with justification for why it works, may be as effective as conceptual instruction. However, in this regard, it is important to note that Perry (1991) found that children who received a combination of conceptual and procedural instruction on mathematical equivalence had transfer performance similar to children who received only procedural instruction, whereas children who received only conceptual instruction had greater transfer performance than either of these two groups. Fourth, children's prior knowledge and educational experiences may influence how they respond to different types of instruction. Children who are frequently encouraged to construct their own procedures or to reflect upon why procedures work may benefit more from procedural instruction than children who receive more traditional instruction. Finally, instruction on multiple procedures, rather than a single procedure, may lead to greater understanding and transfer. Past research has found that students who know multiple procedures have better problem-solving performance and are more likely to learn from instruction than are students who use a single procedure (Alibali & Goldin-Meadow, 1993; Coyle & Bjorklund, 1997; Koedinger & Tabachneck, 1994; Siegler, 1995). Instruction on multiple procedures might be beneficial for conceptual understanding as well.

Future research should address these potential limitations, to assess when conceptual instruction is sufficient on its own, and when and what kind of procedural instruction is appropriate or necessary. To better understand the developmental relations between conceptual and procedural knowledge, future studies should chart the changing relations between the two types of knowledge over the course of weeks or months and should examine the long-term effects of different types of instruction. Additional studies are also needed to extend this work to other domains and to more diverse populations and to explore whether these findings are applicable to more ecologically valid contexts, such as the classroom.

In summary, the present study highlights the causal relations between conceptual and procedural knowledge. Conceptual instruction led to generation of correct, flexible procedures for solving equivalence problems, and procedural instruction led to gains in conceptual understanding. Thus, the relations between conceptual and procedural knowledge are not unidirectional. Instead, conceptual and procedural knowledge appear to develop iteratively, with gains in one type of knowledge leading to gains in the other. However, conceptual knowledge may have a greater influence on procedural knowledge than the reverse. In this study, teaching children the concept behind mathematical equivalence problems, rather than a procedure for solving them, was most effective at promoting flexible problem-solving skill and conceptual understanding. Thus, although there are reciprocal relations between conceptual and procedural knowledge, their influence on one another may not be equivalent.

## References

Alibali, M. W. (1998, June). *The relation between problem encoding and solution strategies: Evidence from mathematical equations.* Poster presented at the 28th Annual Symposium of the Jean Piaget Society, Chicago, Illinois.

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

Alibali, M. W., & Goldin-Meadow, S. (1993). Transitions in learning: What the hands reveal about a child's state of mind. *Cognitive Psychology, 25,* 468–523.

Alibali, M. W., McNeil, N. M., & Perrott, M. A. (1998). What makes children change their minds? Changes in problem encoding lead to changes in strategy selection. In M. A. Gernsbacher & S. Derry (Eds.), *Proceedings of the 20th Annual Conference of the Cognitive Science Society* (pp. 36–41). Mahwah, NJ: Erlbaum.

Baroody, A. J., & Ginsburg, H. (1983). The effects of instruction on children's understanding of the "equals" sign. *Elementary School Journal, 84,* 199–212.

Baroody, A. J., & Gannon, K. E. (1984). The development of the commutativity principle and economical addition strategies. *Cognition and Instruction, 1,* 321–339.

Behr, M., Erlwanger, S., & Nichols, E. (1980). How children view the equal sign. *Mathematics Teaching, 92,* 13–15.

Briars, D., & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology, 20,* 607–618.

Byrnes, J. P. (1992). The conceptual basis of procedural learning. *Cognitive Development, 7,* 235–257.

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

Cauley, K. M. (1988). Construction of logical knowledge: Study of borrowing in subtraction. *Journal of Educational Psychology, 80,* 202–205.

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

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

Cobb, P. (1987). An investigation of young children's academic arithmetic contexts. *Educational Studies in Mathematics, 13,* 109–124.

Cowan, R., & Renton, M. (1996). Do they know what they are doing? Children's use of economical addition strategies and knowledge of commutativity. *Educational Psychology, 16,* 409–422.

Cowan, R., Dowker, A., Christakis, A., & Bailey, S. (1996). Even more precisely assessing children's understanding of the order-irrelevance principle. *Journal of Experimental Child Psychology, 62,* 84–101.

Coyle, T. R., & Bjorklund, D. F. (1997). Age differences in, and consequences of, multiple- and variable-strategy use on a multitrial sort-recall task. *Developmental Psychology, 33,* 372–380.

Dixon, J. A., & Moore, C. F. (1996). The developmental role of intuitive principles in choosing mathematical strategies. *Developmental Psychology, 32,* 241–253.

Frye, D., Braisby, N., Love, J., Maroudas, C., & Nicholls, J. (1989). Young children's understanding of counting and cardinality. *Child Development, 60,* 1158–1171.

Fuson, K. C., & Briars, D. (1990). Using a base-ten blocks learning/teaching approach for first- and second-grade place-value and multidigit addition and subtraction. *Journal for Research in Mathematics Education, 21,* 180–206.

Fuson, K. C. (1988). *Children's counting and concept of number.* New York: Springer-Verlag.

Fuson, K. C. (1990). Conceptual structures for multiunit numbers: Implications for learning and teaching multidigit addition, subtraction, and place value. *Cognition and Instruction, 7,* 343–403.

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.

Gelman, R., & Meck, E. (1986). The notion of principle: The case of counting. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 29–74). Hillsdale, NJ: Erlbaum.

Gelman, R., & Williams, E. M. (1997). Enabling constraints for cognitive development and learning: Domain specificity and epigenesis. In W. Damon (Series Ed.), D. Kuhn, & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., pp. 575–630). New York: Wiley.

Halford, G. S. (1993). *Children's understanding: The development of mental models.* Hillsdale, NJ: Erlbaum.

Hiebert, J., & Wearne, D. (1986). Procedures over concepts: The acquisition of decimal number knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 199–224). Hillsdale, NJ: Erlbaum.

Hiebert, J., & Wearne, D. (1992). Links between teaching and learning place value with understanding in first grade. *Journal for Research in Mathematics Education, 23,* 98–122.

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

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition, 23,* 95–147.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science.* Cambridge, MA: MIT Press.

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

Koedinger, K. R., & Tabachneck, H. J. M. (1994, April). *Two strategies are better than one: Multiple strategy use in word problem solving.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Koedinger, K. R., & MacLaren, B. A. (1997). Implicit strategies and errors in an improved model of early algebra problem solving. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the 19th Annual Conference of the Cognitive Science Society* (pp. 382–387). Hillsdale, NJ: Erlbaum.

Kouba, V. L., Carpenter, T. P., & Swafford, J. O. (1989). Number and operations. In M. M. Lindquist (Ed.), *Results from the 4th Mathematics Assessment of the National Assessment of Educational Progress* (pp. 64–93). Reston, VA: National Council of Teachers of Mathematics.

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

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development, 6,* 449–468.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3,* 359–400.

Pirolli, P., & Recker, M. (1994). Learning strategies and transfer in the domain of programming. *Cognition and Instruction, 12,* 235–275.

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

Resnick, L. B., & Omanson, S. F. (1987). Learning to understand arithmetic. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 3, pp. 41–95). Hillsdale, NJ: Erlbaum.

Rittle-Johnson, B., & Siegler, R. S. (in press). The relations between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.), *The development of mathematical skill.* Hove, UK: Psychology Press.

Sfard, A., & Linchevski, L. (1994). The gain and the pitfalls of reification—The case of algebra. *Educational Studies in Mathematics, 26,* 191–228.

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

Siegler, R. S. (1991). In young children's counting, procedures precede principles. *Educational Psychology Review, 3,* 127–135.

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

Siegler, R. S., & Crowley, K. (1994). Constraints on learning in nonprivileged domains. *Cognitive Psychology, 27,* 194–226.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill.* Cambridge, MA: Harvard University Press.

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36,* 155–193.

Received March 10, 1998
Revision received June 23, 1998
Accepted July 9, 1998 ■