JOURNAL OF COGNITION AND DEVELOPMENT, *6*(2), 285–306

Knowledge Change as a Function of Mathematics Experience: All Contexts are Not Created Equal

Nicole M. McNeil and Martha W. Alibali
University of Wisconsin—Madison

This study investigated how understanding of the equal sign changes as a function of experience in mathematics and variations in context. Students with different levels of mathematics experience (elementary school, seventh grade, undergraduate, and graduate) were randomly assigned to view the equal sign in one of three contexts: (a) equal sign alone, =; (b) typical addition problem, 4 + 8 + 5 + 4 = __; or (c) equivalence problem, 4 + 8 + 5 = 4 + __. Students were asked to define the equal sign and to rate six fictitious students' definitions of the equal sign. Elementary school students interpreted the equal sign as an operational symbol meaning *the answer* or *the total* in all contexts, whereas undergraduate and graduate students viewed it as a relational symbol of equivalence in all contexts. Seventh-grade students interpreted the equal sign as an operational symbol in the alone and addition contexts but as a relational symbol of equivalence in the equivalence context. Results highlight that the shape of knowledge change depends on the context in which the knowledge is elicited. Furthermore, the context may influence whether newly emerging ideas are activated.

Studies of cognitive development are often designed to characterize what people know at different points in developmental time. Studies performed in this tradition imply that individuals either *have* or *lack* knowledge of a particular concept (e.g., children lack knowledge of the conservation of quantity before age 7 but once they reach age 7, they have knowledge of the conservation of quantity). Upon closer inspection, however, it has become increasingly clear that knowledge is more complicated than that. Contemporary theories of knowledge change (e.g., Barsalou, 1982, 1993; Munakata, McClelland, Johnson, & Siegler, 1997; Thelen & Smith,

Requests for reprints should be sent to Nicole McNeil or Martha Alibali at the Department of Psychology, University of Wisconsin-Madison, 1202 West Johnson Street, Madison, Wisconsin 53706. E-mail: nmmcneil@wisc.edu or mwalibali@wisc.edu

1994) suggest that an individual's knowledge of a concept may depend on the context in which the knowledge is elicited. Individuals may understand a concept long before that understanding is reflected in their behavior in a particular context. Equally important, individuals may understand little about a concept but behave as if they do in certain contexts. In this article, we argue that it is essential to consider the context in which knowledge is elicited before drawing conclusions about what people at different experience levels know.

Researchers sometimes overlook the context-dependent nature of knowledge when they are charting how knowledge of a particular concept changes with experience. As a result, there have been gross discrepancies in what knowledge gets attributed to people at different experience levels. One example, which has implications for the way mathematics is taught in schools, has to do with people's knowledge of the equal sign. Some researchers have suggested that children as young as age 11 have a sophisticated interpretation of the equal sign (Perry, Church, & Goldin-Meadow, 1988), whereas others have speculated that even undergraduates have a poor interpretation of the equal sign (Kieran, 1981). In this study, we explore the hypothesis that these discrepant conclusions are due, at least in part, to differences in the contexts in which knowledge of the equal sign has been elicited.

The equal sign is arguably the most fundamental symbol in all of mathematics and science. In mathematics, it often is used to define an equivalence relation (e.g., $3 + 4 = 7$; $x^2 - 9 = [x + 3]\ [x - 3]$), indicating that the expression on the left is the same quantity as the expression on the right. In science, the equal sign is used to express key relationships, such as distance = rate × time, $E = mc^2$, and 1 mole = 6.023×10^{23} atoms. Students must interpret the equal sign as a relational symbol of equivalence if they are to understand certain areas of advanced mathematics and science (e.g., functions).

Despite the importance of the equal sign, traditional kindergarten through 12th-grade American mathematics lessons rarely focus directly on its meaning. Instead, students have to construct an interpretation of the equal sign based on their experiences with it. Not surprisingly, many elementary school students (ages 6–11 years), who have limited experience with mathematics, do not interpret the equal sign in sophisticated ways. Instead of interpreting the equal sign as a relational symbol of equivalence, they tend to interpret it as an operational symbol (Baroody & Ginsburg, 1983; Behr, Erlwanger, & Nichols, 1980; Kieran, 1981; McNeil & Alibali, 2000; Rittle-Johnson & Alibali, 1999; Seo & Ginsburg, 2003). When asked to define the *equal sign*, they often say it means "the total" or "the answer", and when asked to rate the "smartness" of various definitions, they rate definitions such as "the total" or "the answer" as smarter than definitions such as "two amounts are the same" or "equal to" (McNeil & Alibali, 2000, 2002; Rittle-Johnson & Alibali, 1999).

Elementary school students' interpretation of the equal sign parallels their performance solving mathematical equations. Their failure to grasp the relational meaning of the equal sign is particularly salient on equations that have operations on both sides of the equal sign. For example, when presented with the equation 3 + 4 = 5 + 2, most elementary school students say that the equation does not make sense (Behr et al., 1980; see also Baroody & Ginsburg, 1983; Rittle-Johnson & Alibali, 1999). Some students even request that the equation be corrected to 3 + 4 + 5 + 2 = 14, so that it corresponds to their operational interpretation (Behr et al, 1980).

Elementary school students also have great difficulties when they are presented with equations that require them to solve for an operand on the right side of the equal sign (e.g., 3 + 4 + 5 = 3 + __). Equations such as these have been called mathematical equivalence problems in prior work (Alibali, 1999; Perry et al., 1988). When asked to reconstruct mathematical equivalence problems after viewing them for a brief period of time, many students reconstruct the problems as if they were typical addition problems (e.g., they reconstruct 3 + 4 + 5 = 3 + __ as 3 + 4 + 5 + 3 = __; McNeil & Alibali, 2002). This suggests that students think that the equal sign and blank must go together at the end of a problem. When asked to solve equivalence problems, the vast majority of students use incorrect strategies, such as adding up all the numbers in the problem and putting the total amount in the blank (e.g., writing *15* in the blank when solving 3 + 4 + 5 = 3 + __; Perry, et al., 1988; McNeil & Alibali, 2000, 2002). This type of poor performance is consistent with elementary school students' interpretation of the equal sign as the total or the answer.

Although many elementary school students interpret the equal sign as the total or the answer, some more advanced students eventually come to interpret the equal sign as a relational symbol of equivalence. Otherwise, we would not have physicists discovering, understanding, and solving complex equations such as those involving the conservation of energy. We can thus define the boundaries of change in students' thinking about the equal sign. The immature state of thinking is that the equal sign denotes an operation like addition or subtraction and the advanced state of thinking is that the equal sign denotes an equivalence relation between two quantities. Little is known, however, about how and when students' thinking progresses from the immature state to the advanced state.

One widely held hypothesis is that the relational interpretation of the equal sign emerges in middle school, sometime between the ages of 11 and 13. Evidence supporting this hypothesis comes from middle-school students' performance solving equations. Unlike younger students, middle-school students tend to be accurate in their judgments of the truth-value of equations that have operations on both sides of the equal sign (e.g., 3 + 4 = 5 + 2 is true and 3 + 2 = 5 + 3 = 8 is false; Kieran, 1981). Additionally, it is assumed that most middle-school students are able to solve equivalence problems correctly (e.g., 3 + 4 + 5 = 3 + __; Perry et al., 1988). Satisfactory performance solving equations such as these suggests that middle-school students interpret the equal sign as a relational symbol of equivalence.

Contradictory evidence comes from middle-school students' own definitions of the equal sign. When middle-school students are asked to define the equal sign, most provide operational definitions similar to those provided by elementary school students. Very few middle-school students provide relational definitions (Kieran, 1981). Some researchers have speculated that the immature, operational interpretation of the equal sign pervades students' thinking throughout intermediate mathematics, perhaps even up to the level of college calculus (e.g., Kieran, 1981).

We argue that the apparent discrepancies in middle-school students' interpretations of the equal sign can be understood by considering the context in which the equal sign has been presented in past work. In studies of middle-school students' performance solving equations, the equal sign has been presented in the context of an equation that has operations on both sides of the equation (e.g., 3 + 4 = 5 + 2 or 3 + 4 + 5 = 3 + __). In contrast, studies in which students have been asked to define the equal sign have presented the equal sign either alone (=), or in the context of an addition problem (e.g., 3 + 4 = __). We suspect that this context difference across studies is responsible for the observed discrepancies in middle-school students' interpretations of the equal sign. As suggested by prior work (e.g., Munakata et al., 1997; Strohner, 1974), newly emerging ways of thinking about a concept may be particularly sensitive to variations in context.

Barsalou's (1982, 1993) work on concept construction provides a framework for predicting how and when the context can influence thinking. Every concept has a number of ideas, or knowledge chunks, associated with it, and the specific chunks that are activated when the word denoting the concept is perceived can depend on the context. When a chunk is only weakly associated with a concept, the context determines whether the chunk is activated by the concept's word. But when a chunk is strongly associated with a concept, the context plays little role in whether the idea is activated. For example, when individuals hear the word *bird*, they may activate "has feathers" regardless of the context in which they hear the word. This is because the chunk "has feathers" is strongly associated with birds. Indeed, there are not many instances when individuals see a bird without feathers. However, whether individuals activate "can be a pet" when they hear the word *bird* may depend on the context in which the word is presented. This is because the chunk "can be a pet" is only weakly associated with birds. All birds are not pets, and actually, birds are more likely to be seen outdoors than as pets in cages. Thus, if the word *bird* were presented in the context of discussing predatory animals in the wild, it is unlikely that the idea "can be a pet" would be activated. In contrast, if it were presented in the context of discussing dogs and cats, then it is quite likely that "can be a pet" would be activated. Chunks such as "has feathers" are termed context-independent, whereas chunks such as "can be a pet" are termed context-dependent (Barsalou, 1982). The strength of activation of a chunk that is context-dependent varies greatly depending on the context, whereas the strength of a context-independent chunk is relatively strong and highly resistant to contextual manipulations. Moreover, according to Barsalou (1982), a particular chunk can become context-independent after repeated pairings with the word.

We suggest that Barsalou's (1982) work on concept construction provides a useful framework for making predictions about the nature of cognitive change and contextual variability in students' interpretation of the equal sign. In the elementary school years, the equal sign is repeatedly paired with arithmetic operations (see Seo & Ginsburg, 2003, for evidence on this point), so chunks such as "the total" and "the answer" should become strongly associated with the equal sign. Accordingly, the operational interpretation of the equal sign should be activated regardless of the context in which the equal sign is presented.

As children progress through late elementary school and begin middle-school mathematics, they experience the equal sign in the context of equivalent fractions (e.g., 2/3 = __/6), "greater-than, less-than, or equal-to" problems (e.g., is 2 >, <, or = 3?), and pre-algebra problems (e.g., 2 × __ = 6). Thus, chunks such as "the same amount as" and "equal to" should also become associated with the equal sign. Early on, the association between the equal sign and these chunks should be relatively weak compared to that between the equal sign and chunks such as "the total" and "the answer". Accordingly, the context should determine whether chunks such as "the same amount as" and "equal to" are activated when the equal sign is presented.

As students progress through advanced mathematics and science, they encounter problems that reinforce the association between the equal sign and chunks such as "the same amount as" and "equal to". Eventually, "the same amount as" and "equal to" should become so strongly associated with the equal sign that the context should play little role in whether they are activated.

This study tested this account of changes in students' interpretations of the equal sign. We investigated equal sign understanding in students with varying degrees of mathematics experience (elementary school students, seventh-grade students, college undergraduate students who had taken calculus, and physics graduate students). We assessed students' interpretations of the equal sign when presented alone, =; in the context of a typical addition problem, 3 + 4 + 5 + 3 = __; and in the context of an equivalence problem, 3 + 4 + 5 = 3 + __.

Based on our framework, we expected a Mathematics Experience Level × Equal Sign Context interaction. Specifically, elementary school students should interpret the equal sign operationally regardless of context. Seventh-grade students should interpret the equal sign as an operational symbol in the alone and addition contexts, but as a relational symbol of equivalence in the equivalence context. Undergraduate and graduate students should interpret the equal sign relationally regardless of context. Thus, the seventh-grade students' performance should drive the predicted interaction.

METHOD

Participants

Elementary School Students

Fifty-five elementary school students (31 girls, 24 boys; including 20 third graders, 21 fourth graders, and 14 fifth graders) from a religious education program participated. The students attended public elementary schools in a suburban area of Wisconsin. None of the students had had algebra or pre-algebra instruction in their math classes.

Seventh-Grade Students

Twenty-five seventh-grade students (14 girls, 11 boys) from a religious education program participated. The students attended public middle schools in a suburban area of Wisconsin. The students had been introduced to some pre-algebra in their math classes.

Undergraduate Students

Thirty-five undergraduates (24 women, 11 men) from an introductory psychology course at the University of Wisconsin–Madison participated. Students received an extra-credit point for their participation. All students had taken at least one calculus class in their lifetime.

Physics Graduate Students

Twelve physics graduate students (6 women, 6 men) affiliated with the Department of Physics at the University of Wisconsin–Madison participated. All students were beyond their 2nd year of graduate training and had passed the physics qualifying examination, which is an examination that determines retention in the University's doctoral program in physics.

Procedure

Students participated in one experimental session in which they completed a two-page questionnaire that was designed to assess their understanding of the equal sign. The questionnaire was modeled after tasks that have been used in previous work to measure students' conceptions of the equal sign (e.g., Rittle-Johnson & Alibali, 1999; McNeil & Alibali, 2000). A female experimenter administered the questionnaire. Participants were randomly assigned to one of three equal-sign contexts. In the equal sign *alone* context, an equal sign, =, was presented alone at the top of both pages of the questionnaire. In the typical *addition* context, an addition problem, 4 + 8 + 5 + 4 = __, was presented at the top of both pages. In the *equivalence* context, an equivalence problem, 4 + 8 + 5 = 4 + __, was presented at the top of both pages.

The first page of the questionnaire was designed to elicit participants' own definitions of the equal sign. Figure 1 illustrates the equal-sign context manipulation as it was presented on the first page of the questionnaire. Once participants turned to the first page, the experimenter directed the students as follows: "The first question says, 'Tell me what this math symbol means.' There is an arrow on your paper that is pointing to a math symbol, and I want you to tell me what you think that math symbol means. Now, I don't want you to tell me the *name* of the math symbol. I want you to tell me what you think it *means*. You can just write your answer under the question on the black lines." After participants completed the first page, the experimenter instructed them to turn to the second page.

The second page was designed to evoke participants' opinions about a number of possible definitions of the equal sign. At the top of the page, the equal sign was presented in the same context used on page 1. Participants were asked to rate the smartness of six fictitious students' definitions of the equal sign as *very smart*, *kind of smart*, or *not so smart*. The experimenter directed participants as follows: "Some other students told me what they thought that math symbol means. I'm going to tell you what they said, and I want you to circle whether you think it is very

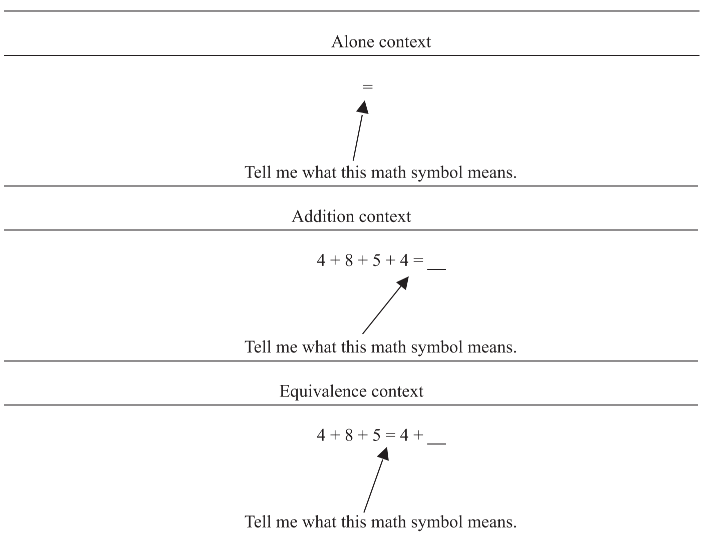

FIGURE 1 Equal-sign context manipulations as they were presented on the first page of the questionnaire.

smart, kind of smart, or not so smart. The first student said that it means *the answer to the problem*. Circle whether you think that is very smart, kind of smart, or not so smart. The second student said that it means *the end of the problem*. The third student said that it means *that two amounts are the same*. The fourth student said that it means *to repeat the numbers*. The fifth student said that it means *that something is equal to another thing*. The last student said that it means *the total*." After participants completed the second page, the questionnaires were collected. The elementary school and seventh-grade students were given a brightly colored pen for participating, and the undergraduates were given one extra credit point.

The physics graduate students completed the two-page questionnaire via electronic mail. Participants were sent the first page of the questionnaire and after they finished and returned the first page, the second page was sent. Instructions for completing the questionnaire were written in the correspondence and were identical to the instructions given to the other experience groups.

Coding Equal-Sign Definitions

Participants' definitions of the equal sign were coded for whether they expressed the operational view of the equal sign (e.g., "the answer," "the total," "add up all the numbers"), the relational view of the equal sign (e.g., "two amounts are the same," "equivalent to"), or an "unspecified equal" view of the equal sign (e.g., "equal," "equals"). Each of the definitions provided by participants in our sample fit into one of the three categories. Examples of participants' definitions are presented in Table 1.

Coding Ratings

Participants' ratings of each fictitious student's definitions of the equal sign were assigned 1 point (*not so smart*), 2 points (*kind of smart*), or 3 points (*very smart*). Students' rating of relational definitions was calculated by summing the ratings for the definitions *two amounts are the same* and *something is equal to another thing*. Students' rating of operational definitions was calculated by summing the ratings for the definitions *the answer to the problem* and *the total*. Students' rating of distracter definitions was calculated by summing the ratings for the definitions *the end of the problem* and *repeat the numbers*.

Reliability of Coding Procedures

Reliability for coding equal-sign definitions was established by having a second coder evaluate the definitions of a randomly selected subsample of approximately 20% of each group (12 elementary students, 5 seventh-grade students, 7

TABLE 1
Example Equal Sign Definitions

Group	*Definition*
Elementary	
Alone	"I think it means equals." [a]
Addition	"It means that the answer is the next thing." [b]
Equivalence	"It means what all the numbers together are." [b]
Seventh grade	
Alone	"When you add, subtract, multiply, or divide, you put this number before your answer." [b]
Addition	"The answer to the question. What all the numbers add up to." [b]
Equivalence	"The right and the left are the same." [c]
Undergraduate	
Alone	"The stuff on the right has the same value as the stuff on the left." [c]
Addition	"Means that it is equal to or equivalent to." [c]
Equivalence	"The numbers on the left equal the numbers on the right." [c]
Graduate	
Alone	"Symbol that compares between quantities/expressions and indicates that they are (or are set to be) equivalent and interchangeable." [c]
Addition	"The sign symbolizes that the numeric value on the left-hand side is the same as the numeric value on the right-hand side." [c]
Equivalence	"It typically represents the concept of equality. For example, = can be spoken in English as 'is equivalent to.'" [c]

[a]Unspecified equal. [b]Operational. [c]Relational.

undergraduates, and 3 physics graduate students). Agreement between coders was 100%.

RESULTS

We predicted an interaction of math experience level and equal-sign context. Specifically, we hypothesized that (a) elementary students would interpret the equal sign operationally regardless of context, (b) seventh-grade students would maintain the operational interpretation in the alone and addition contexts but interpret the equal sign relationally in the equivalence context, and (c) undergraduates and graduate students would interpret the equal sign relationally regardless of context. Because many investigators have found gender differences in mathematics performance (see Hyde, Fennema, & Lamon, 1990, for a meta-analysis), gender was included in all analyses. The main analysis focused on students' own definitions of the equal sign, and a secondary analysis focused on students' smartness ratings of possible definitions of the equal sign. Alpha was set at .05 for all statistical tests.

Students' Own Definitions

We used multinomial logistic regression to predict the log of the odds of giving operational (e.g., *add up all the numbers*), unspecified equal (e.g., *equals*), or relational (e.g., *equivalent to*) definitions. The predictor variables included math experience level (E), equal-sign context (C), and gender (G). The most conservative and easily interpreted way of coding experience level is categorically. Thus, experience level was treated categorically, using three dummy variables to represent its four levels (elementary, seventh grade, undergraduate, and graduate). However, the conclusions were unchanged if experience level was coded using average grade level (4, 7, 14, 18) or years of algebra experience (0, 0.5, 7, 11) as a continuous predictor variable. The graduate student level was used as the reference level. Equal-sign context was treated as a categorical predictor variable, using two dummy variables to represent its three levels (alone, addition, equivalence). The alone context was used as the reference context. Gender was also treated as a categorical predictor variable.

We predicted an interaction of math experience level and equal-sign context. Thus, our hypothesized model included independent effects of math experience level, equal-sign context, and gender, as well as an interaction of math experience level and equal-sign context, that is, E + C + G + E × C. The hypothesized model provided an adequate fit to the data, likelihood-ratio goodness-of-fit statistic = 18.87, $df = 22$, $p = .65$. The hypothesized model was compared to other models using the Likelihood Ratio Test and the Akaike Information Criterion (AIC). The Likelihood Ratio Test compares the –2 log-likelihood of a given model (M_1) to the –2 log-likelihood of a reduced model (M_0) that drops a predictor variable(s) of interest. The difference between the two –2 log-likelihoods, G^2 ($M_0 \mid M_1$), approximates a chi-square distribution with degrees of freedom equal to the difference between the number of parameters in the two models. If G^2 ($M_0 \mid M_1$) is larger than the critical chi-square value, it can be concluded that the log of the odds of falling into one of the three response categories depends on the variable(s) dropped in M_0, and thus, the variable(s) should not be dropped. As shown in Table 2, the log of the odds of giving operational, unspecified equal, or relational definitions of the equal sign were dependent on the interaction of experience level and equal sign context. Thus, the interaction term should not be dropped from the model.

The AIC (Akaike, 1974) was used to ascertain which models were the best models. Essentially, the AIC is a weighted composite of the maximized log-likelihood value for the model and the number of parameters in the model (–2 log-likelihood + 2 [number of parameters]). Models with the smallest AIC are considered optimal. The AIC provides an approximate guide for selecting a model, and small differences in AIC should not be overinterpreted. A good model must have not only a small AIC but also a nonsignificant G^2 ($M_0 \mid M_1$) relative to other potential models, indicating that the log of the odds of falling into one of the response categories is independent of any

TABLE 2
Comparison of Several Models for Estimating the Log of the Odds of Giving an Operational, Unspecified Equal, or Relational Definition of the Equal Sign.

Number	*Model*	*–2logL*	*# par*	*G^2 (M_0 ǀ M_1)*	*AIC*
1	E + C + G	82.67	14	G^2 (1 ǀ 2) = 22.85[a] G^2 (2 ǀ 6) = 3.48	110.67
2	E + C + G + E × C (hypothesized model)	59.82	26	G^2 (2 ǀ 5) = 3.53 G^2 (2 ǀ 4) = 4.98	111.82
3	E + C + G + E × G + C × G	75.31	24	G^2 (3 ǀ 6) = 23.97[a]	123.31
4	E + C + G + E × C + C × G	54.84	30	G^2 (4 ǀ 6) = 3.50	114.84
5	E + C + G + E × C + E × G	56.39	32	G^2 (5 ǀ 6) = 5.05	120.39
6	E + C + G + E × C + E × G + C × G	51.34	36	G^2 (6 ǀ 7) = 10.39	123.34
7	E + C + G + E × C + E × G + C × G + E × C × G	40.95	48	—	136.95

Note. Predictor variables include math experience level (E), equal sign context (C), and gender (G).
[a] Indicates significance; the variable dropped in the model matters and should not be dropped. Thus, dropping the E × C term matters significantly.

variables dropped to achieve the model. As shown in Table 2, both the additive model (E + C + G) and the hypothesized model (E + C + G + E × C) had small AIC values compared to other models. Because the hypothesized model also had a nonsignificant G^2 (M_0 ǀ M_1) when compared to more inclusive models, it can be concluded that the hypothesized model was an optimal model.

The Likelihood Ratio Test presented in Table 2 in which M_1 was E + C + G + E × C (number 2) and M_0 was E + C + G (number 1) indicates that the log of the odds of giving each of the three types of equal-sign definition were dependent on the interaction of mathematics experience level and equal-sign context, G^2 (1 ǀ 2) = 22.85, *df* = 12. Additional evidence for the E × C interaction comes from the Likelihood Ratio Test in which M_1 was the model containing all possible two-way interactions (number 6) and M_0 was the model containing all two-way interactions except E × C (number 3), G^2 (3 ǀ 6) = 23.97, *df* = 12. Thus, the log of the odds of giving operational, unspecified equal, or relational definitions of the equal sign were dependent on the interaction of experience level and equal-sign context. Figure 2 displays the proportion of students at each experience level and equal-sign context in the sample who gave each type of definition. Note the striking effect of context on the definitions given by seventh-grade students.

To examine whether seventh-grade students were more likely to give relational definitions of the equal sign in the equivalence context than in the alone and addition contexts, we constructed 95% confidence intervals (CIs) for the odds ratios comparing the odds of giving relational definitions in the equivalence context to the odds of giving relational definitions in each of the other two contexts. Because

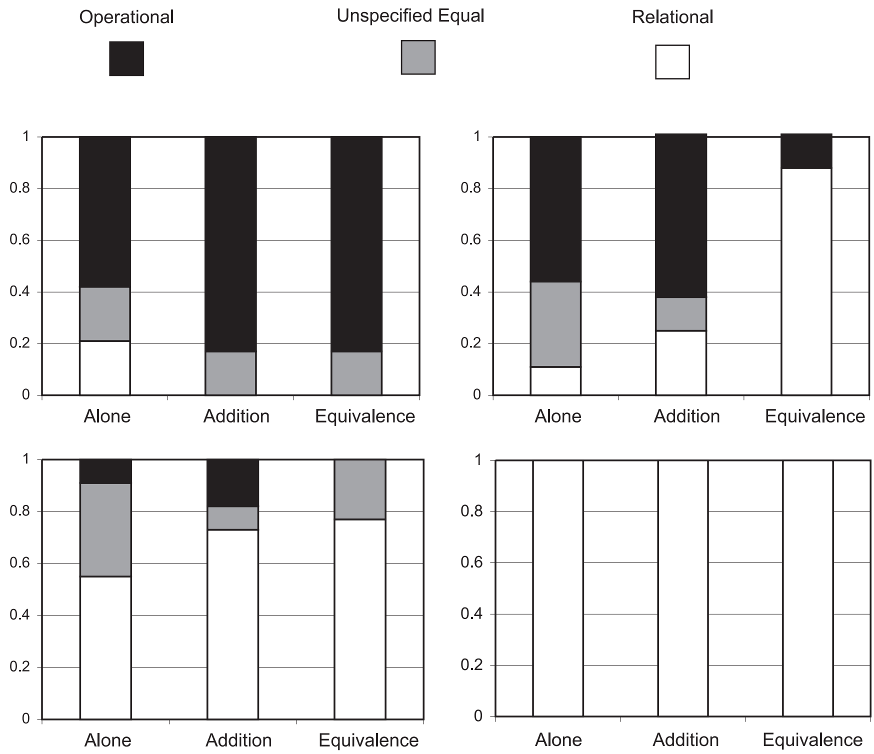

FIGURE 2 Proportion of elementary school students (top left panel), seventh-grade students (top right panel), undergraduates (bottom left panel), and physics graduate students (bottom right panel) in each equal-sign context condition who gave an operational, unspecified equal, and relational definition of the equal sign.

the comparisons involve cell counts that are very small or 0, amended estimators of the odds ratios and asymptotic standard errors were used in which 0.5 was added to each cell count (see Agresti, 1996, p. 25). The odds that seventh-grade students gave relational definitions of the equal sign were 40 times greater in the equivalence context than in the alone context, 95% CI = (4.79, 334.32). The odds that seventh-grade students gave relational definitions of the equal sign were 18 times greater in the equivalence context than in the addition context, 95% CI = (2.06, 157.33). Thus, it can be concluded that the true odds of seventh-grade students giving relational definitions of the equal sign were greater in the equivalence context than in either of the other two contexts.

Likelihood Ratio Tests were also used to examine the independent effects of experience level, equal-sign context, and gender. The Likelihood Ratio Test in which M_1 was E + C + G and M_0 was C + G tested the effect of experience level, controlling for equal-sign context and gender. Not surprisingly, the log of the odds of giving operational, unspecified equal, or relational definitions of the equal sign were dependent on experience level, G^2 (M_0 | M_1) = 78.52, df = 6. As seen in Figure 2, students with less experience were more likely to give operational definitions, and students with more experience were more likely to give relational definitions.

The Likelihood Ratio Test in which M_1 was E + C + G and M_0 was E + G tested the effect of equal-sign context, controlling for experience level and gender. The log of the odds of giving operational, unspecified equal, or relational definitions of the equal sign were not dependent on equal-sign context, G^2 (M_0 | M_1) = 7.06, df = 4.

The Likelihood Ratio Test in which M_1 was E + C + G and M_0 was E + C tested the effect of gender, controlling for experience level and equal-sign context. The log of the odds of giving operational, unspecified equal, or relational definitions of the equal sign were related to gender, but the effect was marginal, G^2 (M_0 | M_1) = 5.06, df = 2, $.05 < p < .10$. Of the 75 female participants, 35 (47%) gave operational definitions, 9 (12%) gave unspecified equal definitions, and 31 (41%) gave relational definitions. Of the 52 male participants, 20 (38.5%) gave operational definitions, 13 (25%) gave unspecified equal definitions, and 19 (36.5%) gave relational definitions. Thus, the differences appears to be in the probability that male participants were more likely than female participants to give unspecified equal definitions of the equal sign. The odds that male participants gave unspecified equal definitions were 2.44 times the odds that female participants gave unspecified equal definitions, 90% CI = (1.32, 4.53). The effect is marginal, so it should be interpreted with caution.

Students' Ratings of Alternative Definitions

We next analyzed students' smartness ratings for operational definitions (*the total*, *the answer*), distracter definitions (*repeat the numbers*, *the end of the problem*), and relational definitions (*equivalent to*, *the same amount as*). For each definition type, students' smartness ratings could range from 2 (*not so smart*) to 6 (*very smart*). Because each participant rated all three types of definitions, the predicted interaction was a three-way interaction of experience level, equal-sign context, and definition type. A 4 (math experience level) × 3 (equal-sign context) × 2 (gender) × 3 (definition type) analysis of variance was performed, with repeated measures on definition type. Mauchly's Test of Sphericity on the variance-covariance matrix indicated that sphericity could not be assumed, approximate χ^2 (df = 2) = 8.4. Consequently, for all tests of within-subjects effects, we multiplied the numerator and denominator degrees of freedom by the relevant Greenhouse–Geisser estimate of epsilon, and for all follow-up comparisons we did not assume homogeneity of variance and used a separate error term for each effect under consideration.

Four effects were significant, and we discuss them in the following order: (a) the predicted three-way interaction of definition type, experience level, and context, $F(11.12, 190.91) = 1.87$; (b) the two-way interaction of definition type and context, $F(3.71, 190.91) = 6.55$; (c) the three-way interaction of definition type, experience level, and gender, $F(5.56, 190.91) = 2.33$; and (d) the main effect of definition type, $F(1.85, 190.91) = 89.76$.

The significant three-way interaction of definition type, experience level, and context is presented in Figure 3. Results support the conclusions of the previous analysis. Seventh-grade students displayed a different pattern of results than did students at the other three experience levels, particularly with respect to the comparisons between the equivalence and alone contexts. Seventh-grade students' ratings of operational definitions were lower in the equivalence context than in the

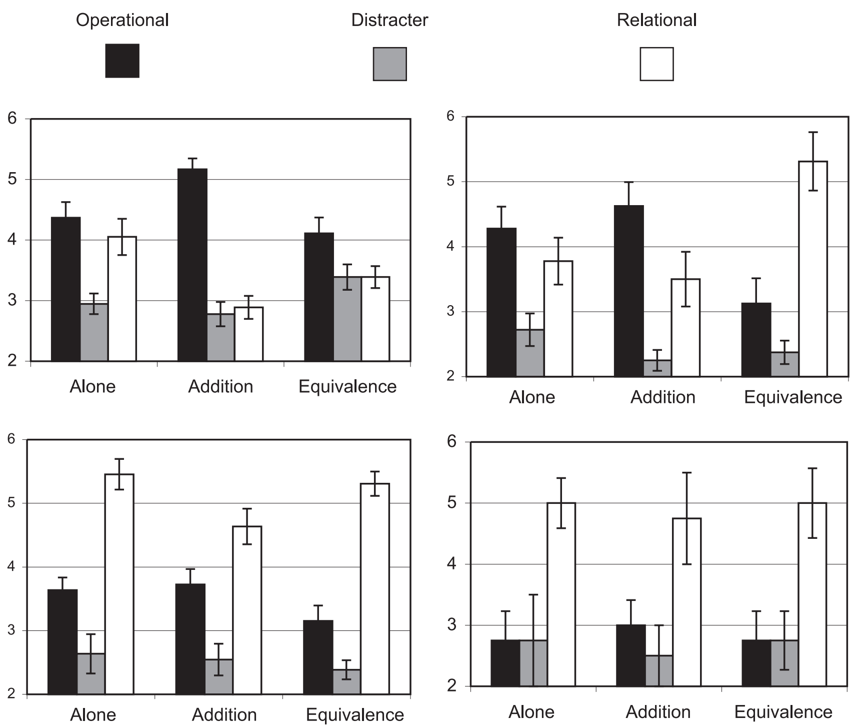

FIGURE 3 Mean *smartness* rating scores given to operational definitions, distracter definitions, and relational definitions by elementary school students (top left panel), seventh-grade students (top right panel), undergraduates (bottom left panel), and physics graduate students (bottom right panel) in each equal-sign context condition. Error bars indicate standard error estimated separately for each group (i.e., not assuming homogeneity of variance).

alone context, $F(1, 22) = 4.78$, and their ratings of relational definitions were higher in the equivalence context than in the alone context, $F(1, 22) = 8.45$.

As can be seen in Figure 3, one of the main differences contributing to the significant three-way interaction was the differential effect of context on seventh-grade students' versus other students' ratings of relational definitions. Whereas the seventh-grade students' ratings of relational definitions differed dramatically between the equivalence context ($M = 5.31$, $SD = 1.28$) and the alone context ($M = 3.78$, $SD = 1.09$), the relational ratings of students in the other experience levels were comparable in the equivalence ($M = 4.29$, $SD = 1.20$) and the alone contexts ($M = 4.62$, $SD = 1.28$). The complex partial interaction comparing the relational ratings of seventh-grade students to those of other students in the equivalence versus the alone context was significant, $F(1, 82) = 7.82$.

The significant two-way interaction of definition type and equal-sign context can be inferred from Figure 3 by averaging across math experience level. Collapsing across experience level, the effect of context differed across definition types, $F(3.15, 195.51) = 6.40$. Overall, context had a significant effect on participants' ratings of relational definitions, $F(2, 126) = 5.49$, and operational definitions, $F(2, 126) = 7.56$, but not on their ratings of distracter definitions, $F(2, 126) = 0.89$. For relational definitions, participants' ratings were lower in the addition context than in the alone context, $F(1, 124) = 10.21$, and ratings were comparable in the equivalence and alone contexts, $F(1, 124) = 0.020$. For operational definitions, participants' ratings were lower in the equivalence context than in the alone context, $F(1, 124) = 4.19$, and they were marginally higher in the addition context than in the alone context, $F(1, 124) = 3.40$, $.05 < p < .10$. Thus, collapsing across experience level, context influenced students' ratings of operational and relational definitions in expected ways but it did not influence students' ratings of distracter definitions.

Figure 4 presents the three-way interaction of definition type, experience level, and gender. Among elementary students, seventh-grade students, and undergraduate students, male and female participants had comparable patterns of ratings across the three definition types: elementary, $F(1.74, 92.23) = 1.17$; seventh-grade, $F(1.25, 28.84) = 0.49$; undergraduate, $F(1.85, 61.01) = 0.45$. However, for graduate students, men and women displayed different patterns of ratings across definition types, $F(1.93, 19.29) = 5.24$. Graduate women rated operational definitions about the same as graduate men, $F(1, 10) = 0.45$; relational definitions marginally higher, $F(1, 10) = 4.62$, $.05 < p < .10$; and distracter definitions lower, $F(1, 10) = 7.27$.

Thus, the main difference contributing to the significant three-way interaction of definition type, math experience level, and gender was the differential effect of gender on graduate students' versus other students' ratings. The complex partial interaction comparing the effect of gender on the pattern of ratings of graduate students to the effect of gender on the pattern of ratings of other students was significant, $F(1.51, 185.98) = 3.08$.

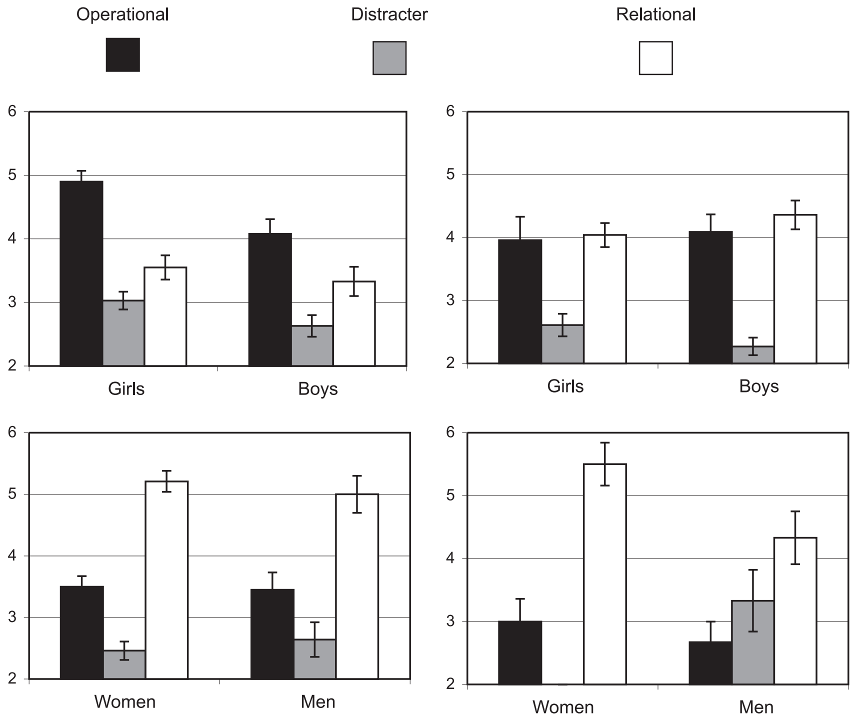

FIGURE 4 Mean *smartness* rating scores given to operational definitions, distracter definitions, and relational definitions by elementary school girls and boys (top left panel), seventh-grade girls and boys (top right panel), undergraduate women and men (bottom left panel), and physics graduate student women and men (bottom right panel). Error bars indicate standard error estimated separately for each group (i.e., not assuming homogeneity of variance).

Finally, the main effect of definition type indicated that overall, students' ratings differed across the three definition types. Students' ratings of distracter definitions ($M = 2.67$, $SD = 0.83$) were lower than their ratings of operational definitions ($M = 3.99$, $SD = 1.18$), $F(1, 126) = 125.14$, and lower than their ratings of relational definitions ($M = 4.20$, $SD = 1.32$), $F(1, 126) = 119.35$. Thus, overall, students viewed distracter definitions as not very smart compared to both operational and relational definitions.

Taken together, results from both analyses support the hypothesis that seventh-grade students' interpretation of the equal sign is highly dependent on context. Seventh-grade students interpreted the equal sign operationally in the alone and addition contexts but relationally in the equivalence context. Across all three contexts, elementary students maintained an operational interpretation and undergraduates and graduate students maintained a relational interpretation.

DISCUSSION

In this study, students' own definitions of the equal sign, as well as their ratings of alternative definitions, provided evidence that the context influences students' interpretations of the equal sign. Most notably, the context influenced the interpretations of seventh-grade students, for whom the relational interpretation is just starting to emerge. As expected, elementary school students defined the equal sign as *the answer* or *the total* regardless of context. Seventh-grade students interpreted the equal sign as *the answer* or *the total* in the alone and addition contexts but interpreted it as a relational symbol of equivalence in the context of an equivalence problem. The undergraduate and graduate students interpreted the equal sign as a relational symbol of equivalence in all contexts, providing evidence that with enough experience, the relational interpretation can supersede the operational interpretation.

The study highlights that scientists and educators need to be cautious when making conclusions about what people know. A widely held belief is that people either have or lack knowledge of a particular concept. Yet, what can be said about seventh-grade students' knowledge of the equal sign as a relational symbol of equivalence? Clearly, their knowledge is not comparable to that of physics graduate students, nor is it comparable to that of elementary school students. When the equal sign was presented alone or in the context of an addition problem, seventh-grade students resembled the elementary school students and appeared to lack knowledge of the equal sign as a relational symbol. However, when the equal sign was presented in the context of an equivalence problem, seventh-grade students resembled the more experienced students and appeared to have knowledge of the equal sign as a relational symbol. Thus, any conclusions about the path of knowledge change in students' interpretation of the equal sign depend on the context in which the knowledge is elicited.

The performance of seventh-grade students suggests that students do not abandon well-established interpretations just because they do not work in a few contexts. Instead, they may view those contexts as exceptions and change their thinking only in those contexts. In middle school, students begin to encounter evidence that contradicts their operational interpretation of the equal sign on a regular basis (e.g., equivalent fractions, pre-algebra problems). At this point, one of three possibilities could occur. First, students could seal themselves off from change completely and continue to maintain their operational interpretation. This possibility is not supported by the data, given that undergraduates and physics graduate students interpreted the equal sign as a relational symbol of equivalence in all contexts. Second, contradictory evidence could provide the impetus for full-scale knowledge reorganization and change. In this case, students' experience with contradictory evidence would lead them to develop a new understanding of the equal sign, thus causing them to shift abruptly from the operational interpretation to the relational interpretation. This possibility also is not supported by the data, given that seventh-grade students maintained the operational interpretation in the alone and addition contexts. Third, students may change their interpretation on a context-by-context basis. That is, contradictory evidence may require only that an exception be made in a particular context, rather than spurring a complete reorganization of knowledge. In this case, students would hold on to their original conception of the equal sign in most contexts, while also interpreting the equal sign as a relational symbol of equivalence in contexts that elicit ideas such as *equivalent to* and *same amount as*. The results of this study support this possibility (see also Seo & Ginsburg, 2003).

The process of making an exception in some contexts, while generally maintaining a different view, has received attention in the prejudice literature (Allport, 1954). Consider an individual who has a negative conception of a particular ethnic group. It is likely that this individual will, at some point, encounter a person from that ethnic group who does not fit the individual's preconception. As the evidence stacks up, it may become so overwhelming that the individual is forced to concede and acknowledge the person as an exception. However, even though the exception is acknowledged, the prejudiced individual will likely hold on to the negative conception of the group as a whole (Allport, 1954). Thus, people sometimes maintain their existing ways of thinking in the face of conflicting environmental input by treating conflicting input as an exception.

If a newly acknowledged exception better accounts for environmental input in a particular context, it will be activated to interpret the input. The exception may be activated in more and more contexts over time and, in this way, may eventually supersede the original concept. Although the original way of thinking is never erased, it may eventually become obsolete, particularly if the exception gets activated more frequently (Siegler, 1999).

In support of this account, the physics graduate students' relational interpretation of the equal sign was relatively immune to context effects. This is not surprising given the nature of the students' extensive experience with the equal sign. As Barsalou (1982) suggested, a particular idea about a concept can become context independent when the idea is repeatedly paired with the concept. For physics graduate students, the bulk of their experience with the equal sign involves problems in which it is essential to interpret the equal sign as a symbol of equivalence. From equations such as "distance = rate × time" that are studied early on in physics training to the computer programming assignment statements such as "int $x = 10$" that graduate students work with on a day-to-day basis, the idea that the equal sign expresses an equivalence relationship is encountered by physicists constantly. After experiencing repeated pairings of the equal sign and the equivalence idea, the physics graduate students develop a firm grasp of the equal sign as a relational symbol and they are able to recognize the relational nature of the equal sign no matter what the context.

This study adds to a growing body of work that addresses gender differences in mathematical abilities. Prior work has suggested that males may have a slight advantage in some areas of mathematics (see Hyde, Fennema, & Lamon, 1990, for a meta-analysis). In this study, gender effects were minor and did not reveal an obvious advantage for either gender. In the analysis of students' own definitions, male participants were more likely than female participants to give unspecified equal definitions of the equal sign. Unspecified equal definitions are less specific than other possible definitions; however, it would be difficult to argue that this implies either an advantage or disadvantage for males, given that students were not directed to be as specific as possible. In students' ratings of alternative definitions, male and female students were comparable at all experience levels, except the graduate level. Graduate student men tended to give distracter definitions higher smartness ratings than did graduate student women. This finding corresponds to prior work that has suggested that gender differences in mathematics are greatest in high-ability samples (Benbow, 1988; Casey, Nuttall, & Pezaris, 1997). However, in this case, giving higher smartness ratings to distracter definitions does not necessarily constitute either an advantage or a disadvantage for males. Thus, this study did not reveal any notable differences between genders in their knowledge of the equal sign.

Our study replicates previous work showing that elementary school students interpret the *equal sign* as "the answer" or "the total" (Baroody & Ginsburg, 1983; Behr et al., 1980; Kieran, 1981). Young students often encounter the equal sign in the context of addition problems, which may be particularly encouraging of the operational interpretation. Consider a problem such as $3 + 4 + 5 + 6 = \_\_$. To solve the problem correctly, children need not interpret the equal sign as a relational symbol of equivalence but rather need only have a strategy for operating on the numbers to get a final answer. This is true not only of typical addition problems but also of most other mathematics problems encountered by elementary school students. Thus, when presented with a mathematics problem, young students may concern themselves with the operations involved in getting the correct solution, and they may come to associate the equal sign with those operations.

If the context consistently reinforces students' arithmetic thinking, as it does year after year in American mathematics classrooms (Valverde & Schmidt, 1997), then students have the potential to become entrenched in the interpretation of the equal sign as *the answer* or *the total*. The ramifications for elementary school students may not be obvious because students who interpret the equal sign as *the answer* or *the total* can perform well in arithmetic, which dominates elementary school mathematics. Indeed, American fourth-grade students have been shown to perform above the international average in mathematics (Mullis et al., 1997).

Although effects may not be immediately apparent, if students hold an entrenched, operational interpretation of the equal sign, this may make the transition to algebra, where the relational meaning of the equal sign is key, particularly difficult (Herscovics & Linchevski, 1994). To truly understand algebraic equations such as $2x + 4 = 18$, students must view the equal sign as a relational symbol of equivalence. Thus, students who are entrenched in an operational interpretation may be at a disadvantage. Consistent with this view, elementary school students who are most entrenched in an operational interpretation of the equal sign are the least likely to benefit from a brief intervention that provides new ways of thinking about equivalence problems (McNeil & Alibali, in press). Thus, although detrimental effects of the operational interpretation of the equal sign may not be apparent in elementary school, they may become marked once students reach algebra. This may be one factor contributing to the below-average performance of American students in international comparisons once they reach eighth grade (Beaton et al., 1996).

Given the central role of context in eliciting newly emerging knowledge, it may be worthwhile for teachers to present the equal sign in a variety of contexts, especially ones that dissuade the operational interpretation, beginning early in mathematics instruction. Additional research is needed to ascertain which combination of contexts would be optimal. Recent work by Seo and Ginsburg (2003) suggests that some methods of exposing elementary school students to different contexts may not be sufficient on their own to foster a relational interpretation in every context. Second-grade students, who were exposed to the equal sign in a variety of contexts in their mathematics class, were tested on their understanding of the equal sign. In class, students had seen the equal sign in a variety of arithmetic contexts (e.g., $2 + 3 = 5$) and nonarithmetic contexts (e.g., 1 dollar = 100 pennies). However, they had not seen it in nontraditional arithmetic contexts (e.g., $5 = 3 + 2$). When students were tested on their understanding of the equal sign in a variety of contexts, they exhibited a relational interpretation of the equal sign only in the nonarithmetic contexts. They maintained the operational view when the equal sign was presented alone and in contexts involving an arithmetic operation, including nontraditional ones such as $5 = 2 + 3$. In contrast, work by Carpenter and colleagues (Carpenter, Franke, & Levi, 2003; Carpenter & Levi, 2000) suggests that even first- and second-grade students have the potential to learn the relational interpretation of the equal sign in arithmetic contexts.

Studies of students' interpretation of the equal sign not only inform mathematics instruction but also contribute to our understanding of cognitive processes more broadly. Research in cognitive development seeks to characterize the knowledge that underlies behavior and to describe the path of knowledge change over time. This study underscores that an individual may appear to have knowledge of a particular concept in one context but not in another. Further, these results highlight that variations across contexts may be particularly important for revealing the nature of newly emerging knowledge. Thus, contextual variation is an important source of information both about how knowledge changes as a function of experience, and about how new knowledge is integrated with old.

ACKNOWLEDGMENTS

This research was supported by a Grant to Martha W. Alibali from the National Science Foundation (BCS-0096129). The research would not have been possible without the support of the administrators, parents, and students at the Sacred Hearts Religious Education Program in Sun Prairie, Wisconsin.

We are grateful to Sabine Lammers and her friends and colleagues in the Department of Physics at the University of Wisconsin. We thank the students and faculty in the Cognitive Development Research Group at the University of Wisconsin for helpful discussions about the study. We also thank Jerry Haeffel, Eric Knuth, Ana Stephens, and two anonymous reviewers for comments on previous versions of the article.

REFERENCES

Agresti, A. (1996). *An introduction to categorical data analysis*. New York: Wiley.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, AC-19,* 716–123.

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

Allport, G. W. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.

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

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10,* 82–93.

Barsalou, L. W. (1993). Challenging assumptions about concepts. *Cognitive Development, 8,* 169–180.

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: Boston College.

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

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioral and Brain Sciences, 11,* 169–232.

Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic and algebra in elementary school*. Portsmouth, NH: Heinemann.

Carpenter, T. P., & Levi, L. (2000). *Developing conceptions of algebraic reasoning in the primary grades*. Madison, WI: National Center for Improving Student Learning and Achievement in Mathematics and Science.

Casey, M. B., Nuttall, R. L., & Pezaris, E. (1997). Mediators of gender differences in mathematics college entrance test scores: A comparison of spatial skills with internalized beliefs and anxieties. *Developmental Psychology, 33,* 669–680.

Herscovics, N., & Linchevski, L. (1994). A cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics, 27,* 59–78.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107,* 139–155.

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

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: Externally imposed goals influence what is learned. *Journal of Educational Psychology, 92,* 734–744.

McNeil, N. M., & Alibali, M. W. (in press). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development*.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1997). *Mathematics achievement in the primary school years: IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: Boston College.

Munakata, Y., McClelland, J. L., Johnson, M. H., & Siegler, R. S. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review, 104,* 686–713.

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

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91,* 175–189.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Seo, K.-H., & Ginsburg, H. P. (2003). "You've got to carefully read the math sentence … ": Classroom context and children's interpretations of the equals sign. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills* (pp. 161–187). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Siegler, R. S. (1981). Developmental sequences within and between concepts. *Society for Research in Child Development Monographs, 46,* (Whole No. 189).

Siegler, R. S. (1999). Strategic development. *Trends in Cognitive Science, 3*, 430–435.

Strohner, H., & Nelson, K. (1974). The young child's development of sentence comprehension: Influence of probability, nonverbal context, syntactic form and strategies. *Child Development, 45,* 567–576.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

Valverde, G. A., & Schmidt, W. H. (1997/1998, Winter). Refocusing U.S. math and science education. *Issues in Science and Technology,* 60–66.