Cognitive Science 32 (2008) 366–397
Copyright © 2008 Cognitive Science Society, Inc. All rights reserved.
ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1080/03640210701863933

Trade-Offs Between Grounded and Abstract Representations: Evidence From Algebra Problem Solving

Kenneth R. Koedinger[a], Martha W. Alibali[b], Mitchell J. Nathan[c]

[a]*Human–Computer Interaction Institute, Carnegie Mellon University*
[b]*Department of Psychology, University of Wisconsin–Madison*
[c]*Department of Educational Psychology, University of Wisconsin–Madison*

Received 21 August 2005; received in revised form 10 February 2007; accepted 16 February 2007

Abstract

This article explores the complementary strengths and weaknesses of grounded and abstract representations in the domain of early algebra. Abstract representations, such as algebraic symbols, are concise and easy to manipulate but are distanced from any physical referents. Grounded representations, such as verbal descriptions of situations, are more concrete and familiar, and they are more similar to physical objects and everyday experience. The complementary computational characteristics of grounded and abstract representations lead to trade-offs in problem-solving performance. In prior research with high school students solving relatively simple problems, Koedinger and Nathan (2004) demonstrated performance benefits of grounded representations over abstract representations—students were better at solving simple story problems than the analogous equations. This article extends this prior work to examine both simple and more complex problems in two samples of college students. On complex problems with two references to the unknown, a "symbolic advantage" emerged, such that students were better at solving equations than analogous story problems. Furthermore, the previously observed "verbal advantage" on simple problems was replicated. We thus provide empirical support for a trade-off between grounded, verbal representations, which show advantages on simpler problems, and abstract, symbolic representations, which show advantages on more complex problems.

Keywords: Problem solving; External representation; Algebra; Grounding; Abstraction

1. Introduction

External problem representations have a profound effect on problem-solving performance and learning (Collins & Ferguson, 1993; Day, 1988; Kirshner, 1989; Zhang, 1997). As one example, different external representations of the Tower of Hanoi problem lead to different rates of correct solution (Kotovsky, Hayes, & Simon, 1985). However, understanding of how specific characteristics of external representations influence performance and learning

Correspondence should be addressed to Ken Koedinger, Human–Computer Interaction Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213. E-mail: koedinger@cmu.edu

is limited. One dimension along which representations vary is in how grounded or abstract they are (Paivio, 1986; Palmer, 1978). In this article, we focus on how grounded and abstract representations influence performance for problems of varying complexity and for participants of various levels of competency. Our theoretical analysis suggests that there is a fundamental trade-off between grounded and abstract representations in supporting problem solving. Specifically, we hypothesize that grounded representations are more effective than abstract representations for simpler problems like those typically encountered early in learning, whereas abstract representations are more effective for more complex problems like those encountered later in learning. If true, this trade-off has important implications for the use and ordering of alternative representations, both in performance environments (e.g., when are features of the Mac vs. Unix operating system better for users) and in learning environments (e.g., should mathematics story problems be presented before or after formal equation-solving exercises).

We investigate the hypothesized trade-off in the context of algebra problem solving. Algebra is a natural choice because, apart from natural language, it is the first abstract symbolic language that most people learn (i.e., it is usually learned before other abstract symbolic languages such as programming languages, chemical equations, and so forth). Algebraic reasoning is also important in high school and post-secondary education. In prior research, Koedinger and Nathan (2004) found that students solved story problems more successfully than matched equations. Story problems were often solved without using abstract symbolic equations, by using informal strategies that are grounded in concrete knowledge of quantitative relations. In other words, for such problems, students experience a verbal advantage whereby they are better able to solve problems presented in verbal form than in corresponding symbolic form. If Koedinger and Nathan's results generalize to all algebra story problems, that is, if algebra story problems are generally easier than corresponding equations, one may justifiably wonder why we teach equation solving at all. We suspected, however, that such a generalization may not be true for the entire range of algebra problems, and we set out to identify whether there are some problems, namely complex problems, for which equations are easier than corresponding story problems.

For more complex problems, we hypothesize that a "symbolic advantage" emerges, such that using the abstract language of equations enhances performance relative to reasoning with grounded story representations. In this article, we explore the evidence for this representation-complexity trade-off and consider its general implications for the study of problem solving and learning.

Our analysis is focused on external representations, which can be written on paper, though it may also apply to internal mental representations. We define grounded representations as ones that are more concrete and specific, in the sense that they refer to physical objects and everyday events. Abstract representations, in contrast, leave out any direct indication of the physical objects and events they refer to and hence are more general as well as more concise.

In the context of quantitative reasoning, real-world problem situations or "story problems" are more grounded than symbolic equations because they use familiar words and refer to familiar objects and events.[1] For example, consider the following story problem.

> Ted works as a waiter. He worked 6 hours in one day and also got $66 in tips. If he made $81.90 that day, how much per hour does Ted make?

Given some experience with money and waiters, the words, objects, and events in this problem are relatively familiar. Students' understanding of the quantitative relationships described is thus grounded in familiar terms and in knowledge of corresponding objects and events.

In contrast, consider the following symbolic representation of the story problem above.

$$x^{*}6 + 66 = 81.90$$

The equation form is clearly shorter or more concise than the story problem (12 characters as compared to 108). The equation leaves out any reference to familiar objects and events like hourly wages and tips. Furthermore, the terminology is different. The semantics of the "words" in the algebraic sentence (i.e., x, *, =) are likely to be less familiar to beginning algebra students than the phrases expressing analogous meanings in the story (i.e., how much, hours in one day, he made). Although others have identified student difficulties with algebraic symbols such as variables (Clement, 1982; Knuth, Alibali, McNeil, Weinberg, & Stephens, 2005; Küchemann, 1978) and the equal sign (Kieran, 1981; McNeil & Alibali, 2004; Rittle-Johnson & Alibali, 1999), the difficulty of symbolic expressions relative to matched verbal expressions had not been explored before Koedinger and Nathan (2004).

2. Trade-offs in representational advantages

Some previous studies have demonstrated benefits for grounded or concrete representations (e.g., Koedinger & Nathan, 2004; Nunes, Schliemann, & Carraher, 1993; Paivio, Clark, & Khan, 1988). Others have demonstrated benefits for abstract representations (Day, 1988; S. H. Schwartz, 1971, 1972; Sloutsky, Kaminski, & Heckler, 2005). We present a framework for reasoning about the circumstances under which different representations are most effective.

We explore different properties of grounded and abstract representations and how these can yield different computational advantages and disadvantages (see Table 1). Grounded representations tend to be more familiar, so their meanings can be more readily accessed in long-term memory (Table 1, first row). For example, in the story problem above, the words (e.g., "tips," "gets") and the syntactic structure (e.g., "also got [number] in tips," "how much per [unit] does [person] make") are familiar to young adults. However, to correctly comprehend the equations, students must remember the meanings of less familiar formal "words" (e.g., x, =) and syntax (e.g., order of operations), which may be more difficult to access in long-term memory. Koedinger and Nathan (2004) observed greater student difficulty on equations than stories like those above, and based on student errors, they provided an explanation in terms of lack of familiarity with the "words" and syntax of equations. Beginning algebra students have greater prior experience with aspects of the English language used to express quantitative relationships than with the abstract language of algebra.[2]

Besides being more familiar, grounded representations tend to be more reliable, in the sense that students are less likely to make errors and more likely to detect and correct them when they are made (Table 1, second row). This reliability is a consequence of redundant semantic elaborations that are connected with grounded representations and that can be used to support or check inferences (cf. Baranes, Perry, & Stigler, 1989; Hall, Kibler, Wenger, & Truxaw, 1989; Nhouyvanisvong, 1999). Abstract representations are stripped of such semantic

Table 1
Trade-offs in computational characteristics of more grounded versus more abstract representations

		Type of Representation and Level of Benefit		
Benefit	By Means of Property	Grounded	⇔	Abstract
Ease of LTM Access	Familiarity	Higher		Lower
Reliability	Redundancy	Higher		Lower
WM support	Externalizability	Lower		Higher
Efficiency	Conciseness	Lower		Higher
	Examples:	Stories		Equations

Note. LTM = Long-term memory; WM = Working memory.

elaborations; thus, errors are more likely to be made and more likely to go unnoticed. For example, for the Waiter story problem presented above, Koedinger and Nathan (2004) found that students never added the number of hours worked to the amount of tips. However, in the corresponding equation, students frequently added $6 + 66$. Such undetected errors in equation solving are fairly common, even for equation-solving experts (C. H. Lewis, 1981).

Although abstract representations may be more error prone than grounded representations, abstract representations have several advantages. Working with abstract representations can be fast and efficient because their concise form allows for quick reading, manipulating, and writing (Table 1, fourth row). Consequently, abstract representations put fewer demands on working memory than grounded representations because it is easier to use paper as an external memory aid (Table 1, third row). Further, one need not keep track of the referents of all the symbols while solving the problem, and one may more easily mentally imagine manipulations of quantitative relations (e.g., combination of like terms) when using abstract representations (cf. Kirshner, 1989).

3. Representational advantages in the acquisition of algebra skill

Contrary to common belief and prior claims in the literature (e.g., Cummins, Kintsch, Reusser, & Weimer, 1988; Geary, 1994), Koedinger and Nathan (2004) found that high school students succeeded more often on grounded, story problems than on matched abstract equations. This finding is consistent with other results showing that problem situations can activate real-world knowledge and aid problem solution (Baranes et al., 1989; Carraher, Carraher, & Schliemann, 1987; Hall et al., 1989; Hudson, 1983).

Koedinger and Nathan (2004) provided a two-part explanation for this pattern. First, students were less successful on symbolic equations than one might have expected. Students' errors in the symbolic format often revealed serious difficulties with the syntax and semantics of equations. Students made errors in comprehending and manipulating algebraic expressions (see also Matz, 1980; Payne & Squibb, 1990; Sleeman, 1986). For example, students often violated syntactic rules such as order of operations or performed illegal algebraic manipulations (e.g., subtracting from both sides of the plus sign rather than the equal sign). Second, students were more successful on story problems than one might have expected. In contrast with normative expectations, students often did not solve the story problems by converting them to equations and then manipulating symbols. Instead, students usually used informal strategies, not involving algebraic symbols, to "bootstrap" their way to correct answers. For example, students sometimes used an iterative guess-and-test procedure to arrive at a solution, and they sometimes worked backwards through the constraints provided in the problem, "unwinding" the constraints. We refer to these as informal strategies because they do not involve the domain formalism (algebra symbols in this case) and are not traditionally formally taught in school.

A detailed cognitive model of these results, the Early Algebra Problem Solving (EAPS) theory, was developed by Koedinger and MacLaren (2002). The EAPS theory contains comprehension processes, represented as ACT-R production rules (Anderson & Lebiere, 1998), that describe how students process the given external problem representation and create internal representations of the quantitative structure. The EAPS model accurately predicts student error patterns and frequencies. A key to this prediction is the idea that students' equation comprehension skills are worse than their verbal comprehension skills; that is, they can comprehend word problems, but have trouble comprehending equations.

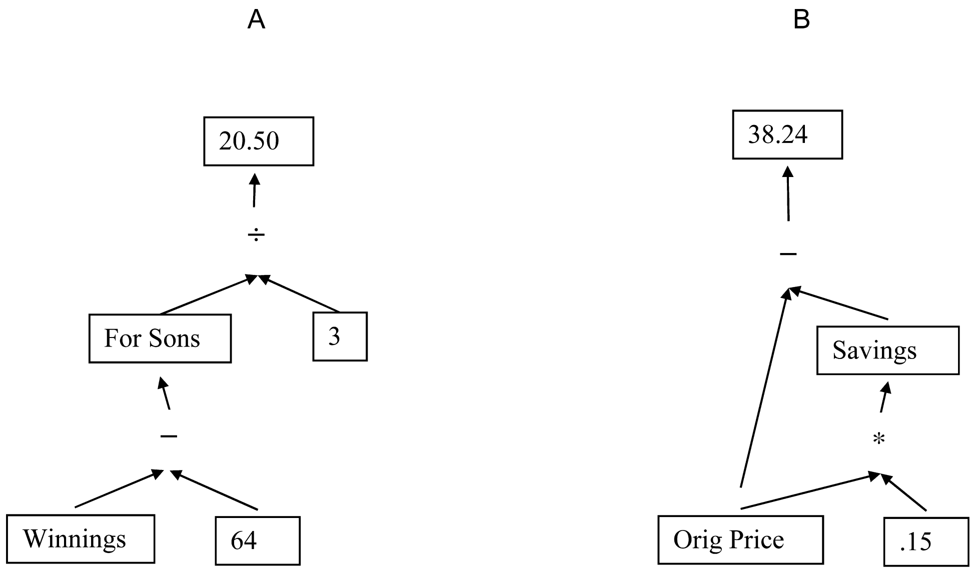
Figure 1a shows the network of quantitative relations for the Lottery story problem. The EAPS cognitive model can solve such problems using various strategies, including the informal unwind strategy mentioned above. To unwind, EAPS searches for a quantitative relationship for which the output is known and all but one of the inputs are known (an English version of the corresponding production rule is shown in Fig. 1c). In the Lottery problem, the second quantitative relation (the division node) satisfies the condition ("if" part) of this production—the amount each son got is known (20.50) and the number of sons is known (3) and thus the portion for all sons can be computed.

Students are able to use informal strategies like "unwind" to solve problems without recourse to algebra equations. We lack both an empirical and theoretical base for knowing when and how abstract formalisms, like algebra equations, might be superior to such informal strategies, and given past results (Koedinger & Nathan, 2004; Nhouyvanisvong, 1999), we should not simply assume that they are.

In this article we explore the hypothesis that the advantages of grounded representations hold true for simple problems, but the advantages of abstract representations emerge for more complex problems. Early in the acquisition of a formal skill, students can succeed with grounded representations by using informal strategies that do not require abstract formalisms. They fail with abstract representations because they have difficulty comprehending them. To the beginner, formalisms like algebra are like a foreign language (see Ernest, 1987). For more complex problems that are presented later in skill acquisition, this pattern may reverse. First, students increasingly acquire familiarity and facility with abstract formalisms. Second, as problems become more complex, limitations of informal strategies emerge and become increasingly more severe.

For example, the unwind strategy depends on being able to invert operators and work backwards toward the problem unknown. However, the unwind strategy is thwarted when the unknown is referenced more than once. For instance, consider the following double-reference problem:

> Roseanne just paid $38.24 for new jeans. She got them at a 15% discount. What was the original price?



Mom won some money in a lottery. She kept $64 for herself and gave each of her 3 sons an equal portion of the rest of it. If each son got $20.50, how much did Mom win?

(X - 64) ÷ 3 = 20.50

Roseanne just paid $38.24 for new jeans. She got them at a 15% discount. What was the original price?

X - 0.15X = 3824

C

Production rule to "unwind" quantitative relations:

IF the output of a quantitative relation is known
 and one input is known and the other input is unknown
THEN
 Set subgoals to find the inverse of the operator
 And to apply it to the known output and input to find the unknown input

Fig. 1. Double-reference problems are more complex than single-reference problems because they thwart the unwind strategy that works on single-reference problems. (a) Examples of analogous single-reference story and equation problems and their underlying quantitative structure. (b) Examples of analogous double-reference story and equation problems and their underlying quantitative structure. (c) An English version of the unwind production rule in the EAPS cognitive model.

We call such problems *double-reference* problems because the unknown (the original price) is used or referenced twice in the quantitative relationships described in the problem. The original price is referenced once in a multiplication relation (the original price *times* 15% is the savings) and once in a subtraction relation (the original price *minus* the savings is the amount paid). The key structural differences between single- and double-reference problems are evident in the EAPS theory's "quantitative network" for this Discount problem (Fig. 1b) as compared to the Lottery problem (Fig. 1a). In contrast to the Lottery problem, in the Discount problem neither quantitative relation satisfies the condition of the unwind production (Fig. 1c). In neither relation is there a known output and input. In the first relation, only one input is known, the discount rate of 15%, and the output (savings) and other input (original price) are both unknown. In the second relation, both inputs are unknown.

In double-reference problems, one cannot work backwards from a known result to the unknown because the unknown is in more than one place (see also Bednarz & Janvier, 1996).[3] Without an effective informal strategy for solving a complex double-reference problem, students can turn to the algebraic strategy of translating to an equation and solving that equation. However, given students' well-documented difficulties with translation (Heffernan & Koedinger, 1997; Mayer, 1982; Nathan, Kintsch, & Young, 1992), this strategy is prone to error.

In this article, we present two experiments that test for a representation-complexity trade-off. These experiments were designed to test whether the verbal advantage previously observed on simpler problems would extend to more advanced students and to demonstrate, for the first time, a symbolic advantage for more complex problems. These experiments also show, for the first time, a representation-complexity trade-off within the same sample of students.

Experiment 1 investigates this trade-off for students with fairly weak mathematics skills, that is, college students in an algebra review course. Such students are expected to have limited success with complex problems. Experiment 2 investigates whether this trade-off holds for highly skilled students, who are expected to have greater success with complex problems. In the following sections, we first present the two experiments and the main results regarding the representation-complexity trade-off. We then provide a detailed analysis of students' performance and errors in an effort to explain the sources of the observed trade-off.

4. Experiment 1

4.1. Method

4.1.1. Participants

Participants were 153 students from two college-level algebra courses at a state university. Forty-three were in a basic elementary algebra course (similar to high school Algebra 1) and 110 were in a more advanced intermediate algebra course (content combines high school Algebra 1 and 2 topics).

4.1.2. Procedure

Students were given 20 minutes within class time to complete a six-item "difficulty factors assessment," which systematically varies hypothesized factors that may affect problem difficulty. Each student was randomly assigned one of six test forms. The forms were designed to contrast the two difficulty factor dimensions, representation and problem complexity, illustrated in Table 2. We manipulated problem representation, creating mathematically equivalent problems in both story and equation representations. Two types of story problems

Table 2
Six problem categories illustrating two difficulty factors used in Experiment 1: representation and number of unknown references

Representation	Number of Unknown References: Single	Number of Unknown References: Double
Story-implicit operators	Mom won some money in a lottery. She kept \$64 for herself and gave each of her 3 sons an equal portion of the rest of it. If each son got \$20.50, how much did Mom win?	Roseanne just paid \$38.24 for new jeans. She got them at a 15% discount. What was the original price?
Story-explicit operators	After hearing that Mom won a lottery prize, Bill took the amount she won and subtracted the \$64 that Mom kept for herself. Then he divided the remaining money among her 3 sons giving each \$20.50. How much did Mom win?	Roseanne bought some jeans on sale for \$38.24. To figure that sales price, the salesperson took the original price, multiplied it by the 15% discount rate, and then subtracted the outcome from the original price. What was the original price?
Equation	Solve for the unknown value, X: $(X - 64) \div 3 = 20.50$	Solve for the unknown value, X: $X - 0.15X = 38.24$

were used: (a) more narrative and familiar story-implicit problems in which arithmetic operations are expressed implicitly through everyday verbs (e.g., "kept") and (b) more equation-like story-explicit problems in which arithmetic operations are explicitly expressed (e.g., "subtracted"). Examples are presented in the first two rows of Table 2. This manipulation addresses a concern that the story-explicit problems used in Koedinger and Nathan (2004) may not be representative of story problems as they occur in the real world and in textbooks; story-implicit problems are more representative. The third row in Table 2 illustrates the equation representation.

The columns in Table 2 illustrate the problem complexity dimension. The simpler single-reference problems, illustrated in the first column, are a subset of the single-reference problems used by Koedinger and Nathan (2004). The more complex double-reference problems are illustrated in the second column. Although there are other complexity differences between the single- and double-reference problems besides the number of references to the unknown (e.g., cover story and base equation differences), the key hypothesis being tested here is that story problems will be easier than equations within the single-reference problems, and harder than equations within the double-reference problems.

For each level of problem complexity, there were three different cover stories, each of which involved a different base equation. Combining these six cover stories with the three representation types yielded 18 different problems. These problems were distributed onto six forms such that: (a) each problem appeared on two different forms with its position counterbalanced (i.e., the first position on one form, sixth on the other) and (b) all six cover stories appeared on each form, with each of the three representations appearing at both levels of complexity.

Table 3
Strategy codes and coding definitions

Strategy	Definition
Algebra	Student uses algebraic manipulations to derive solution
Arithmetic	Student works forward to derive solution using the operations presented in the problem (applies primarily to the result-unknown problems used in Experiment 2)
Unwind	Student works backward to derive solution using inverse operations
Untangle	Student informally combines problem constraints and then works backward to derive solution using inverse operations
Guess and test	Student generates and tests potential solutions
Answer only	Student provides solution without showing any written work

4.1.3. Coding

Each problem was scored as correct or incorrect. Next, the strategy used to solve the problem was coded, using the scheme outlined in Table 3. For problems that were scored as incorrect, error type was coded, using the scheme outlined in Table 4.

Reliability was established by having a second trained coder rescore the data for 15 participants. Agreement between coders was 96% ($N = 90$) for problem-solving strategies. For error types, agreement between coders was 93% ($N = 46$) using the three broad categories in Table 4 (no response, conceptual error, arithmetic error) and 84% ($N = 46$) using the finer-grained categories in Table 4.

4.2. Results and discussion

As seen in Fig. 2, on the simpler single-reference problems, students performed better with the more grounded story representations than with the more abstract equation representation.

Table 4
Error codes and definitions (broad categories used in Fig. 6 are in italics)

Error Type	Definition
No response	Student leaves problem blank
Conceptual errors	
Give up	Student performs some work but does not provide a final answer
Order of operations	Student violates order of operations rule
Comprehension	Student shows evidence of either incorrect interpretation of problem constraint(s) or a failure to produce external forms (equations or arithmetic) consistent with those constraints
Missing operator	Student does not perform one of the operations presented in the problem
Bad algebra	Student performs an incorrect algebraic manipulation, such as subtracting from both sides of the plus sign rather than both sides of the equal sign
Inversion	Student fails to invert an operator (e.g., change + to −) that needs to be inverted to solve the problem
Answer only	Student writes only a number and the number is the wrong
Arithmetic error	Student adds, subtracts, multiplies, or divides wrong
Copy slip	Student miscopies a value given in the problem or previously generated

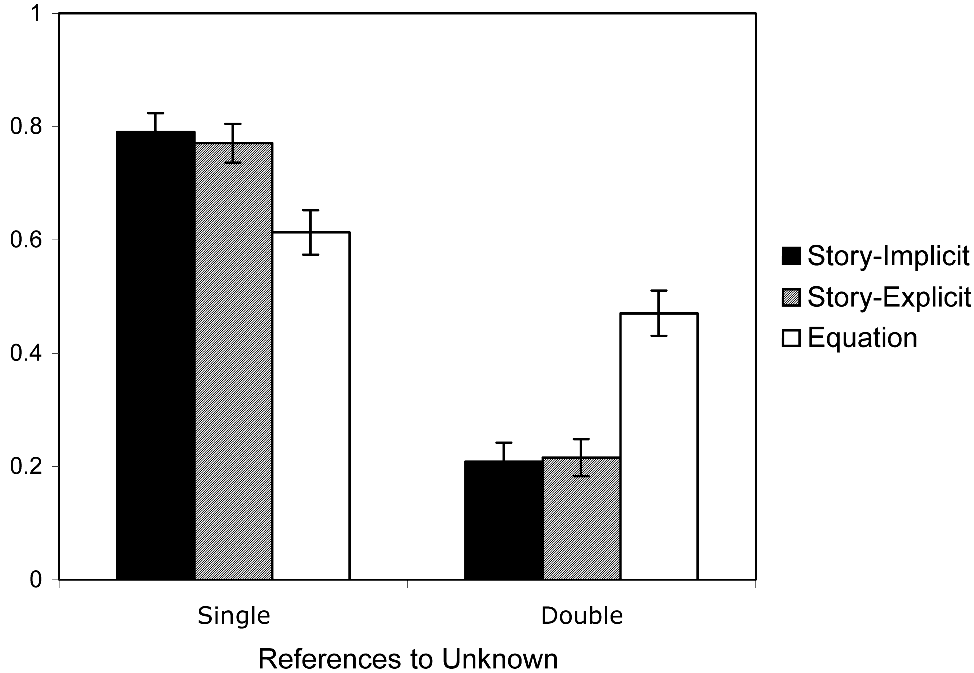


Fig. 2. The interaction of problem complexity (number of references to the unknown) and representation in Experiment 1. Error bars represent standard errors.

In contrast, on the more complex double-reference problems, students performed better with the abstract equation representation than with the story representations. This interaction was significant, $F(2, 304) = 27.46$, $p < .0001$. An item analysis (using item, not subject, as the random effect; see Clark, 1973) suggests that this interaction effect may also generalize across items, $F(2, 8) = 3.29$, $p < .10$, though more than six core items would be needed for a more conclusive result. There was also a main effect of problem complexity, $F(1, 152) = 246.36$, $p < .0001$. There were no differences between story-implicit and story-explicit problems, on either single- (79 vs. 77%) or double-reference problems (21 vs. 22%).

The finding that these college students succeed more often on single-reference stories than on single-reference equations not only replicates the verbal advantage previously observed among high school students (Koedinger & Nathan, 2004) but also generalizes it in two ways. First, our use of story-implicit problems demonstrates that the verbal advantage is not limited to the story-explicit problems used by Koedinger and Nathan. Second, this study demonstrates that the verbal advantage applies to college algebra students who perform better overall than the high school algebra students in Koedinger and Nathan's study (72 vs. 49% correct on single-reference problems, respectively).

The results for individual problems are presented in Table 5. Looking at the single-reference stories and equations (first three rows), we see that, for the most part, these college students have mastered two-operator, single-reference problems. However, whereas their story-problem–solving competence is uniformly high (averaging 80, 71, and 83% on the three problem types), their equation-solving competence is spotty. For the two problems in the familiar

Table 5
Problems used and proportion of participants who solved each problem correctly in Experiment 1

Story Problems		Equations	
Problem	Proportion of Participants who Solved Correctly	Problem	Proportion of Participants who Solved Correctly
Single-Reference Problems			
Laura bought 7 donuts and paid $0.12 extra for the box to hold them. If she paid $2.57 total, what is the price per donut?	0.80	$7X + .12 = 2.57$	0.79
Ted works as a waiter. He worked 6 hours in one day and also got $46 in tips. If he made $65.50 that day, how much per hour does Ted make?	0.71	$6X + 46 = 65.50$	0.80
Mom won some money in a lottery. She kept $64 for herself and gave each of her 3 sons an equal portion of the rest of it. If each son got $20.50, how much did Mom win?	0.83	$(X - 64) \div 3 = 20.50$	0.23
Double-Reference Problems			
There are 38 students in class. If there are 6 more girls than boys, how many boys are in the class?	0.54	$X + (X + 6) = 38$	0.71
Roseanne just paid $38.24 for new jeans. She got them at a 15% discount. What was the original price?	0.04	$X - 0.15X = 38.24$	0.29
You are in Paris, France, and you want to exchange your dollars for French Francs (FF). The first exchange store gives you 5.7 FF per dollar but charges 22 FF for each exchange. The second exchange store gives you 5.4 FF per dollar and does not charge a fee. When are the charges from the two stores the same? In other words, what amount of dollars results in the same charge from both stores?	0.05	$5.7X - 22 = 5.4X$	0.41

Note. The story problems shown are the implicit operator versions, but the data are the mean proportion correct for both versions.

"mx + b = y" form, students performed just as well on the equations (79%) as on the stories (76%). However, for the equation in the "(x − c) / n = y" form, which involves less familiar elements (parentheses and "/"), students performed much worse (23%). This performance difference is striking, and it is consistent with the notion that students are acquiring the language of algebra in pieces. These students have acquired pieces of algebra language knowledge for an equation form that is highly frequent in textbooks but have not fully acquired the pieces needed for another form that is rare in textbooks.

The results for the double-reference problems are quite different. For all three cover stories and their corresponding equations (Table 5, last three rows), students performed better on the equations than on the story problems. Thus, we see that introducing the complexity of multiple references to the unknown leads to an advantage for symbols.

Strictly speaking, the representation-complexity trade-off hypothesis is about the representations that students use, not the representations they are given. The analysis presented thus far is an indirect test of this hypothesis. It is based on the assumption that the given external representation influences the representation and corresponding strategy the problem solver uses. This assumption seems reasonable given the preponderance of evidence that external representations profoundly affect performance (e.g., Collins & Ferguson, 1993; Kirshner, 1989; Kotovsky et al., 1985; Zhang, 1997). Nevertheless, students have been observed to use a variety of representations and corresponding strategies when given story problems (diSessa, Hammer, Sherin, & Kolpakowski, 1991; Greeno & Hall, 1997; Hall et al., 1989). Therefore, we also address the representation-complexity trade-off hypothesis more directly by examining the strategies students used to solve the problems.

We grouped students' solution strategies into two broad categories: abstract, formal strategies (algebra) and more grounded, informal strategies (arithmetic, unwind, untangle, guess-and-test, and answer only; see Table 3 for definitions). Solutions that involved multiple strategies, one of which was algebra, were classified with formal strategies. This analysis is summarized in the "% Used" columns of Table 6. As expected, students in Experiment 1 rarely used informal strategies on equations (1% of single-reference, 1% of double-reference equations), but they often did so on stories (55% of single-reference, 27% of double-reference stories). Thus, the given representation clearly influenced students' strategy choices.

The key question is whether students are more effective with formal strategies on the more complex story problems and informal strategies on the less complex story problems.

Table 6
Percentage of problems solved and solved correctly with informal and formal strategies for story problems and equations

	Story Problems				Equations			
	Informal		Formal		Informal		Formal	
	% Used	% Correct	% Used	% Correct	% Used	% Correct	% Used	% Correct
	Experiment 1							
Single-reference (start unknown)	55	86	42	73	1	50	94	65
Double-reference	27	7	52	37	1	0	90	52
	Experiment 2							
Single-reference (result unknown)	91	97	9	100	68	89	32	95
Single-reference (start unknown)	32	86	68	98	11	86	88	88
Double-reference	12	58	88	80	4	71	96	88

This analysis is summarized in the "% Correct" columns of Table 6. Indeed, the predicted interaction holds. On the simpler single-reference stories (row 1, left columns), students were more effective when they chose more grounded, informal strategies (86% correct) than when they chose more abstract, formal strategies (73% correct). Conversely, on the more complex double-reference stories (row 2, left columns), students were more effective when they chose formal strategies (37% correct) than when they chose informal strategies (7% correct).

Such between-subject comparisons may suffer from selection bias; that is, the strategies may not be better, but instead better students may be selecting these strategies. One way to address this concern is to focus on individuals who used an informal strategy on one problem and a formal strategy on another problem of the same type. On single-reference stories, 36 students used a formal strategy on one problem and an informal strategy on another. These students were more successful on single-reference stories when they used the informal strategy (92% correct) than when they used the formal strategy (75% correct; $F[1, 35] = 3.9$, $p = .057$). On double-reference stories, 46 students used an informal strategy on one problem and a formal strategy on another. On these problems, they were more successful when they used the formal strategy (37% correct) than when they used the informal strategy (7% correct; $F[1, 45] = 12.2$, $p < 0.01$). Thus, we not only see a trade-off between grounded and abstract representations based on the representation given to students, we also see the same trade-off based on the external (written) representations students use in solving problems.

We next present a replication of these results with a more mathematically sophisticated student sample, and we then proceed to a detailed analysis of students' errors in both experiments.

5. Experiment 2

Mathematically more sophisticated students should be more adept at comprehending and manipulating abstract formal representations. Compared to the students in Experiment 1, mathematically sophisticated students have developed a greater familiarity with abstract symbolic representations (Table 1, row 1) and thus may have little difficulty accessing the knowledge needed to comprehend abstract representations. In other words, the hypothesized "ease of long-term memory access" benefit of grounded representations should have limited relevance for such students.

Given greater facility with abstract representations, students should be more likely to employ them as an aid in problem solving. However, just because a student uses an abstract external representation does not mean that a grounded interpretation of that representation is lost (cf. Hall et al., 1989). Although these students may tend to use equations even on simple story problems, they may solve these equations with the redundant semantic support of the given grounded representation. In other words, they may reap the hypothesized reliability benefit of grounded representations (see row 2 of Table 1) and perform better on these simpler story problems than on analogous context-free equations, even though it may appear from their written work that they are using the abstract representation alone. Thus, we predict that because of the reliability benefit of grounded representations, even these highly experienced students will perform better on stories than equations when in single-reference form.

In brief, Experiment 2 both attempts to replicate the representation-complexity trade-off observed in the first experiment and to explore whether we can separate the effects of the two different hypothesized benefits of grounded representations. We do so by using a sample that should show effects of reliability, but limited effect of ease of long-term memory access.

5.1. Method

5.1.1. Participants

Sixty-five undergraduate students from Carnegie Mellon University participated in partial fulfillment of a course requirement. Students' self-reported Math SAT scores had a mean of 719 (range 510–800).

5.1.2. Procedure

Students completed a difficulty factors assessment designed to evaluate equation solving and story problem solving. The assessment also included two other types of items that will not be considered here: drawing generalizations from tables and translating between different representational formats.

All problem-solving items were presented in both story and equation format. All story problems were of the familiar, implicit-operator type. As shown in Table 7, the single-reference problems included both those where the unknown is the result of the process described, "result unknowns" (e.g., $800 - 30 * 23 = y$), and those where the unknown is the start of the process described, "start unknowns" (e.g., $600 - 20 * x = 260$). The double-reference problems were identical to those used in Experiment 1.

5.1.3. Coding

Each problem was scored as correct or incorrect. Next, the strategy used to solve the problem was coded using the scheme outlined in Table 3, and for problems that were scored as incorrect, error type was coded using the scheme outlined in Table 4.

Reliability for problem-solving strategies was established by having a second trained coder rescore the data for 10 participants. Agreement between coders was 96% ($N = 100$). Because this sample contained few incorrectly solved problems, an additional set of 30 incorrectly solved problems was sampled to check error coding. Agreement between coders was 93% ($N = 30$) using the three broad categories in Table 4 (no response, conceptual error, arithmetic error) and 85% ($N = 30$) using the finer-grained categories in Table 4.

5.2. Results and discussion

As seen in Fig. 3, the predicted representation-complexity trade-off also held in this sample, $F(1, 64) = 17.89$, $p < .0001$. Students showed a verbal advantage for the single-reference problems, $F(1, 64) = 5.82$, $p < .02$, as well as a symbolic advantage for the double-reference problems, $F(1, 64) = 12.74$, $p < .001$. Results for individual problems are presented in Table 7. For both types of single-reference problems, students showed a verbal advantage, and for all three types of double-reference problems, students showed a symbolic advantage. An item analysis (using item, not subject, as the random effect, see Clark, 1973) suggests that this

Table 7
Problems used and proportion of participants who solved each problem correctly in Experiment 2

Story Problems		Equations	
Problem	Proportion of Participants Who Solved Correctly	Problem	Proportion of Participants Who Solved Correctly
Single-Reference Problems			
Anne is in a rowboat on a lake. She is 800 yards from the shore. She rows toward the shore at a speed of 30 yards per minute. How far is Anne from the shore after 23 minutes?	0.97	$800 - 30 * 23 = y$	0.91
Kim is saving up for a mountain bike that costs $600. She earns $20 per week by babysitting every Saturday afternoon. She can save all of the money that she earns. If Kim only needs to save $260 more, how many weeks has she already saved?	0.94	$600 - 20 * x = 260$	0.88
Double-Reference Problems			
There are 38 students in class. If there are 6 more girls than boys, how many boys are in the class?	0.86	$X + (X + 6) = 38$	0.92
Roseanne just paid $38.24 for new jeans. She got them at a 15% discount. What was the original price?	0.63	$X - 0.15X = 38.24$	0.75
You are in Paris, France, and you want to exchange your dollars for French Francs (FF). The first exchange store gives you 5.7 FF per dollar but charges 22 FF for each exchange. The second exchange store gives you 5.4 FF per dollar, and does not charge a fee. When are the charges from the two stores the same? In other words, what amount of dollars results in the same charge from both stores?	0.82	$5.7X - 22 = 5.4X$	0.94

Note. The numbers shown in this table illustrate one of the two number sets used in the experiment. Each student saw a story problem using one number set and a matched equation using the other.

interaction effect may also generalize to new items, $F(1, 3) = 38.4$, $p < .01$. Because we used some new items (the first two in Table 7) that were not in Experiment 1 or Koedinger and Nathan (2004), the item generality of the representation-complexity interaction is further strengthened, though a larger sampling of single- and double-reference problems would be desirable.

As in Experiment 1, we also examined not only the representations students were given but also the representations they actually used to solve the problems (Table 6, last three rows). The given representation clearly influenced students' choice of solution strategies.[4] Students rarely used informal strategies on equations (7 of 64 start-unknown single-reference equations, 11%;

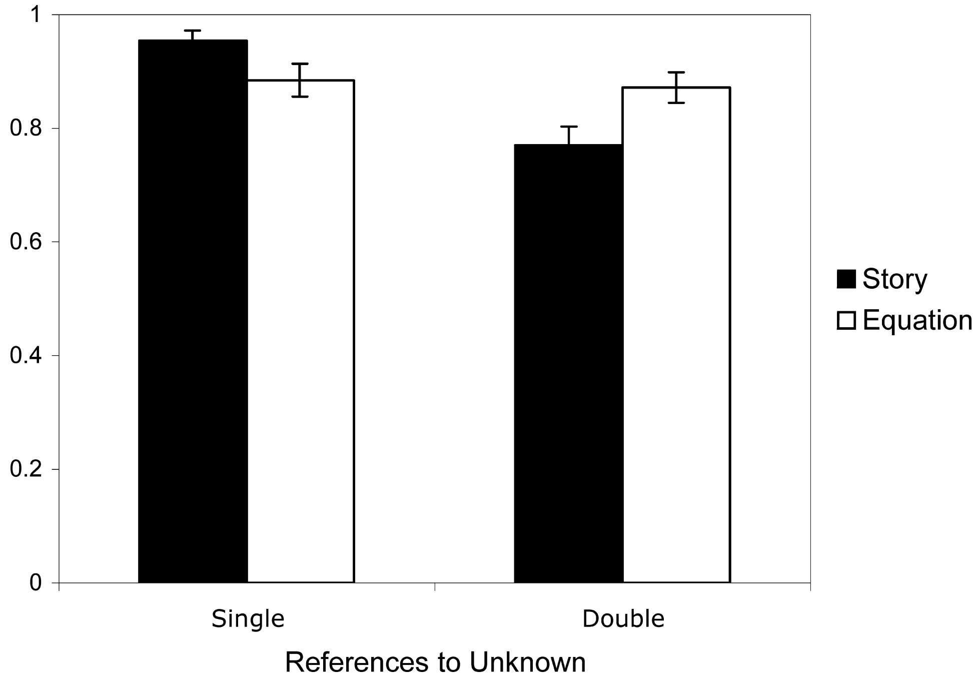


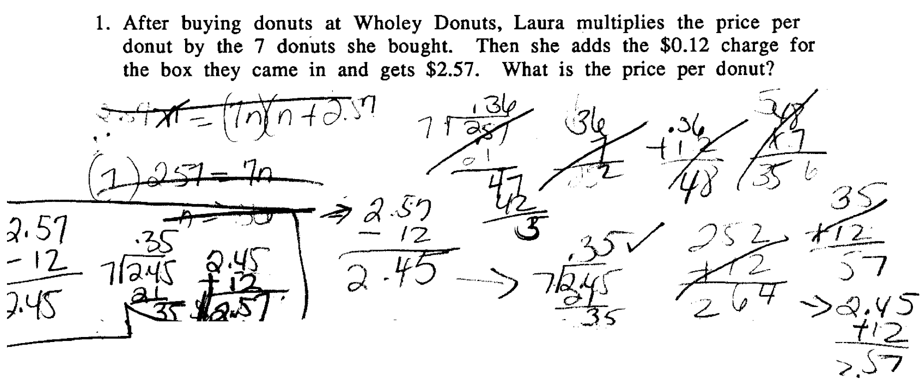
Fig. 3. The interaction of problem complexity (number of references to the unknown) and representation in Experiment 2. Error bars represent standard errors.

7 of 195 double-reference equations, 3.6%), but they did so about three times more often on the corresponding stories (21 of 65 start-unknown single-reference stories, 32%; and 24 of 195 double-reference stories, 12%). As in Experiment 1, students in Experiment 2 were more successful on double-reference stories when they chose formal strategies (80%) than when they chose informal strategies (58%). However, unlike Experiment 1, this formal advantage also held for the simpler story problems: students in Experiment 2 were more effective on start-unknown single-reference stories when they chose formal strategies (98%) than when they chose informal strategies (86%). Interestingly, this strategy performance difference provides evidence for the reliability benefit of grounded representations. Despite having little or no difficulty with retrieving the semantics of equation forms, these students still perform better on start-unknown stories than matched equations. Unlike students in Experiment 1, who do so primarily because they use informal strategies as an alternative to equation solving, the more advanced students in Experiment 2 are primarily using equations on start-unknown story problems. However, they more reliably solve those equations in the context of a grounded story (98%) than in abstract isolation (88%).

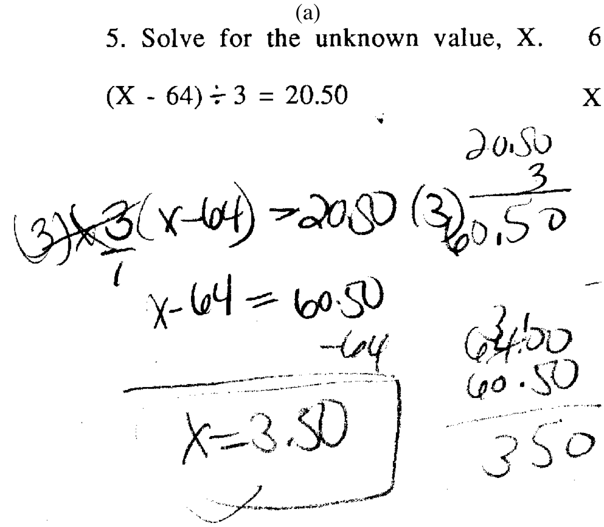
Given their success with formal strategies, it may seem surprising that the sophisticated students in this experiment used informal strategies at all. One possible explanation is that students sometimes perceived informal strategies to require less effort (cf. Scribner, 1986). Indeed, for single-reference problems, the informal unwind strategy requires fewer production rules (mental steps) in the EAPS cognitive model described above and involves less written work because no equations are needed, just arithmetic (see Fig. 4c vs. 4b).

6. Sources of the representation-complexity trade-off

Toward a causal explanation of the representation-complexity trade-off, we turn next to a detailed analysis of students' performance and errors. Across both experiments, students

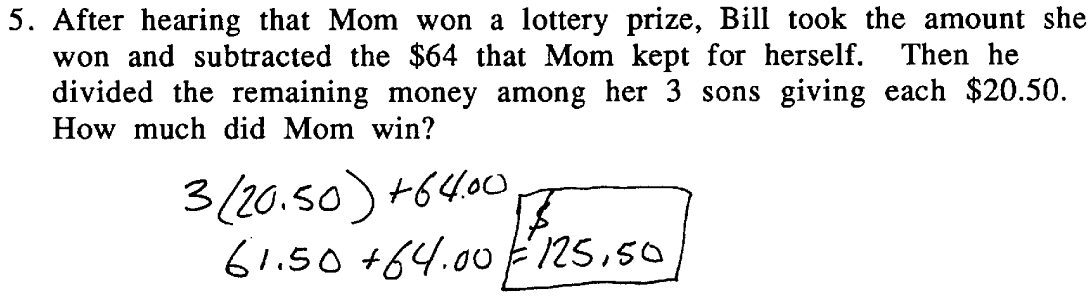


(a)

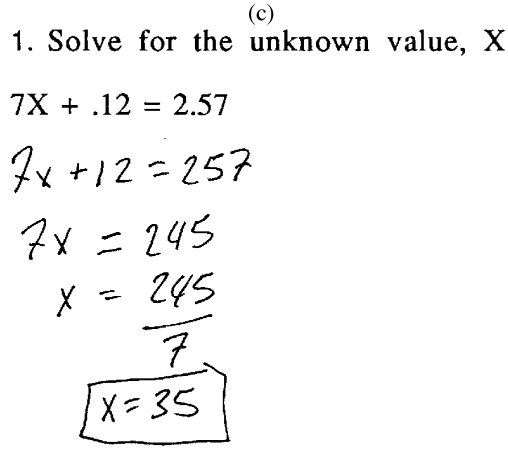


(b)

Fig. 4. Student work illustrating the verbal advantage on single-reference problems. (a) Student 87 successfully using unwind on the Donuts single-reference story problem. (b) Student 87 (same as in a) incorrectly using equation solving on the single-reference Lottery equation. The conceptual error is in the second step where the student subtracts 64 rather than adding 64 to both sides. (c) Student 46 successfully using unwind on the Lottery single-reference story problem. Note indications of money semantics in the addition of "$ " in the final answer and of ".00" to the given value of "64." Such extra semantic relations coming from the grounded story may explain the contrast between the student's success here and the same student's error on an abstract equation, shown just below (d). (d) Student 46 (same as in c) incorrectly using equation solving on the single-reference Donuts equation. The error is in the first step where the student is "multiplying through" to convert decimal terms, .12 and 2.57, to whole numbers but does not know or retrieve the semantic requirement that the conversion must be applied to all terms, including converting 7x to 700x. (Continued)



(c)



(d)

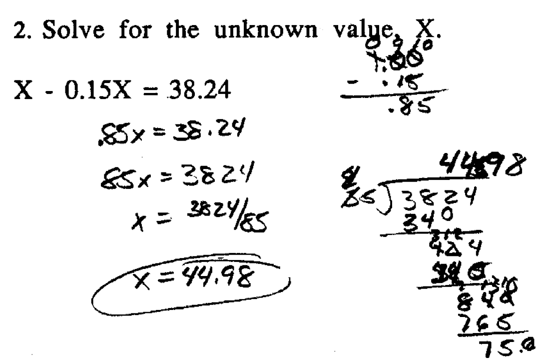
Fig. 4. (*Continued*)

succeeded more often on single-reference problems when the problems were presented in grounded story form than when they were presented as abstract equations. This verbal advantage on simple problems is illustrated with examples of student work in Fig. 4. Each student succeeded on a story version of a single-reference problem (donuts for student 87, lottery for student 46) but failed on an equation version of these same problems (lottery for student 87, donuts for student 46). Both students correctly solved each story problem without using formal algebra. Instead, each successfully applied the informal unwind strategy, exploiting grounded relations. Consider the contrast between the error in Fig. 4b of subtracting 64 rather than adding it when given an abstract equation and the success in Fig. 4c when given a grounded story. The student in 4c has the benefit of the grounded relations that the amount Mom won is made up of, and thus must be more than, the $64 she kept for herself.

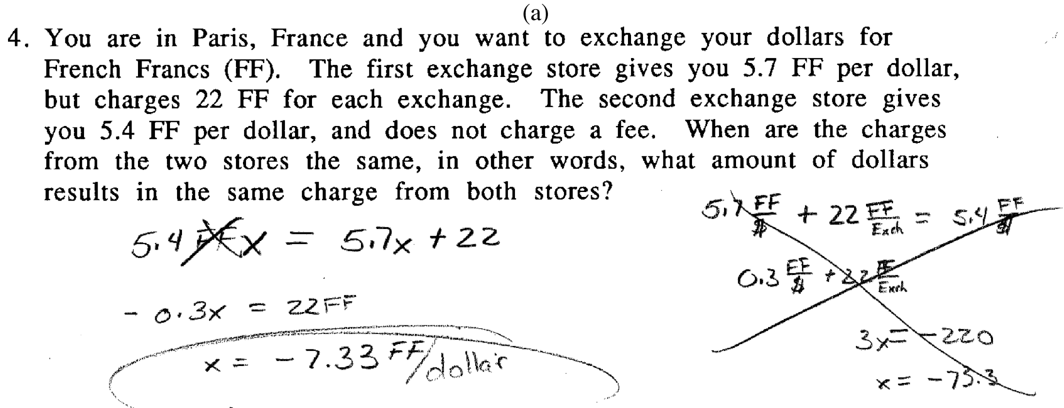
In addition to finding a verbal advantage for simple problems, we also found a symbolic advantage for complex problems in both experiments. Representative examples of student work are shown in Fig. 5. Both students succeeded on an equation version of a double-reference problem (discount for student 19, exchange for student 82), whereas each failed on a story version of these problems (exchange for student 19, discount for student 82). Both students attempted to translate the story problem into an equation but made errors in this symbolization process. However, both succeeded in solving equivalent equations.

What are the sources of the verbal advantage for single-reference problems and the symbolic advantage for double-reference problems? To better understand the patterns illustrated in Figs. 4 and 5, we examined the nature of students' errors and compared these with predictions from our analysis of the computational characteristics of grounded and abstract representations (Table 1).

Following Koedinger and Nathan (2004), we categorized student errors into three broad categories: (a) no response, (b) arithmetic error, and (c) conceptual errors. A solution was

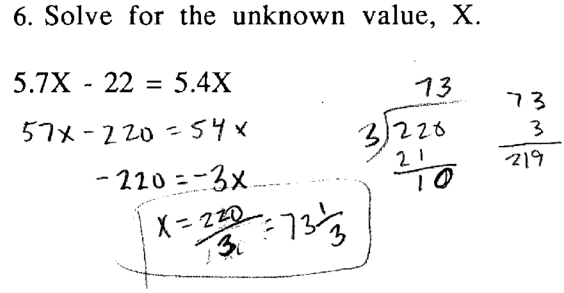


(a)

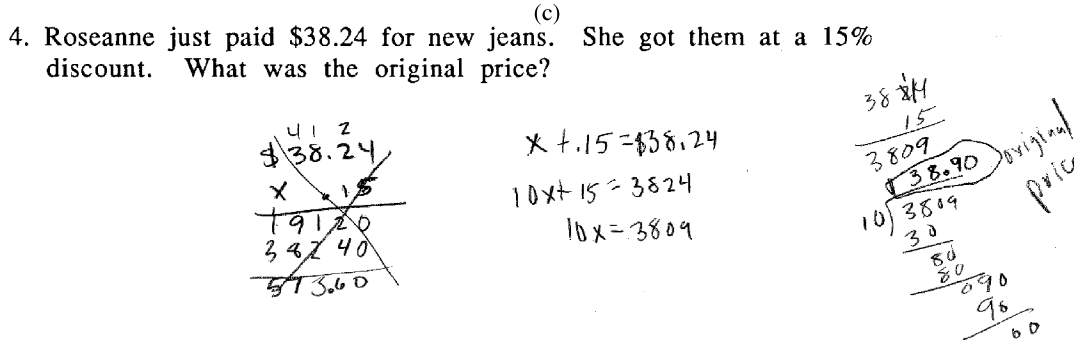


(b)

Fig. 5. Student work illustrating the symbolic advantage on double-reference problems. (a) Student 19 successfully solves the double-reference equation for the discount problem. (b) Student 19 fails on the double-reference story problem (Exchange story). The student performs an incorrect translation to an equation (the "+ 22" should be added to 5.4x, not 5.7x) and then does not appear to notice that the negative result of equation-solving (−7.33) is an unlikely answer to this problem situation. (c) Student 82 successfully solves the double-reference equation for the Exchange problem. (d) Student 82 fails on the Discount story problem. The student attempts both an incorrect informal strategy (multiplying 38.24 by.15 on the left) and an incorrect translation to an equation ("x + .15" should be "x − .15x"). (Continued)



(c)



(d)

Fig. 5. (*Continued*)

coded as No Response if nothing was written down for that problem. A solution was coded as an Arithmetic Error if the solution was conceptually correct; that is, all required arithmetic operations are written down correctly but one of these arithmetic operations was incorrectly performed. Apart from some rare nonarithmetic "copy slips" such as incorrectly copying a digit from the problem statement (1.5% of solutions in Experiment 1 and 0.5% of solutions in Experiment 2), all other errors were coded as Conceptual Errors. See Table 4. (Note that answer-only errors were extremely rare, only 0.7% or 6 of 918 solutions, and answers given were not consistent with other observed arithmetic errors; such errors were categorized as conceptual errors to indicate greater severity.)

6.1. Sources of the verbal advantage on single-reference problems

The familiarity property of grounded representations (Table 1, row 1) makes it more likely that grounded knowledge is retrieved from long-term memory when needed. If students fail to retrieve relevant knowledge, they may give up and not respond. Thus, for single-reference problems, we expect to see more No Response errors on equations than on story problems, particularly for students with less algebra background.

The redundancy property of grounded representations (Table 1, row 2) makes it more likely that grounded knowledge is applied reliably. Students can use the semantic cues available in grounded representations to facilitate reliable performance on arithmetic operations (Baranes et al., 1989; Nunes et al., 1993). For instance, Koedinger and Nathan (2004) found that students were less likely to make place-value alignment errors (e.g., add 15.90 and 66 and get

16.56) when these numbers are grounded as money quantities (they do not add 66 *dollars* to 90 *cents*). Because equations do not provide such semantic redundancy, we expect to see more Arithmetic Errors on equations than on story problems. We predicted that this reliability benefit of grounded representations would facilitate arithmetic performance even for the well-prepared students in Experiment 2.

While No Response errors indicate low familiarity and Arithmetic Errors indicate low reliability, Conceptual Errors may result from either. A solution with a conceptual error involves at least some response, but it has at least one error more severe than an arithmetic error.

We predict a shift in the kinds of errors underlying the verbal advantage on single-reference problems. With increasing algebra skill, we should see a shift from a predominance of familiarity-related errors (mostly No Response) to a mix of familiarity- and reliability-related errors (mostly Conceptual) and eventually to a predominance of reliability-related errors (mostly Arithmetic) on equations. This pattern is just what we see across three student populations of increasing competence for single-reference problems: (a) high school students' difficulties with equations are reflected mostly in No Response errors (Koedinger & Nathan, 2004), (b) less prepared college students' equation difficulties are reflected mostly in Conceptual Errors (Experiment 1, Fig. 6, top), and (c) highly prepared college students' equation difficulties are reflected mostly in Arithmetic Errors (Experiment 2, Fig. 6, bottom).

For students in Experiment 1 with intermediate algebra experience (Fig. 6, top), the verbal advantage is driven primarily by more Conceptual Errors on equations (21%) than on story problems (8%). Unlike the high school students in Koedinger and Nathan (2004), for whom the verbal advantage was driven primarily by more No Response errors on equations, these college-level algebra students have developed enough familiarity with symbolic equations to at least attempt most problems. However, their still-limited familiarity with equations may lead to Conceptual Errors, like subtracting from both sides of the plus sign (not equal sign) or multiplying through to get rid of decimals but ignoring the whole numbers (see Fig. 4d). Students may also make reliability-related Conceptual Errors (e.g., Inversion in Table 4) while performing algebraic operations on equations without the redundancy support of the grounded story representation.

For students in Experiment 2 with lots of algebra experience (Fig. 6, bottom), the verbal advantage is driven primarily by more Arithmetic Errors on equations (6.2%) than on story problems (1.5%). Even for students who are very familiar with formal symbols (note the complete absence of No Response errors), the redundancy benefit of the grounded story representation helps them avoid or correct arithmetic errors that they might make in solving ungrounded equations.

6.2. Sources of the symbolic advantage on double-reference problems

What factors led to the symbolic advantage observed on double-reference problems? In predicting errors on the double-reference problems, we considered the positive characteristics of abstract representations (Table 1, first two columns), but focus here on working memory support. Because we did not collect solution-time data, we cannot directly address the issue of efficiency.

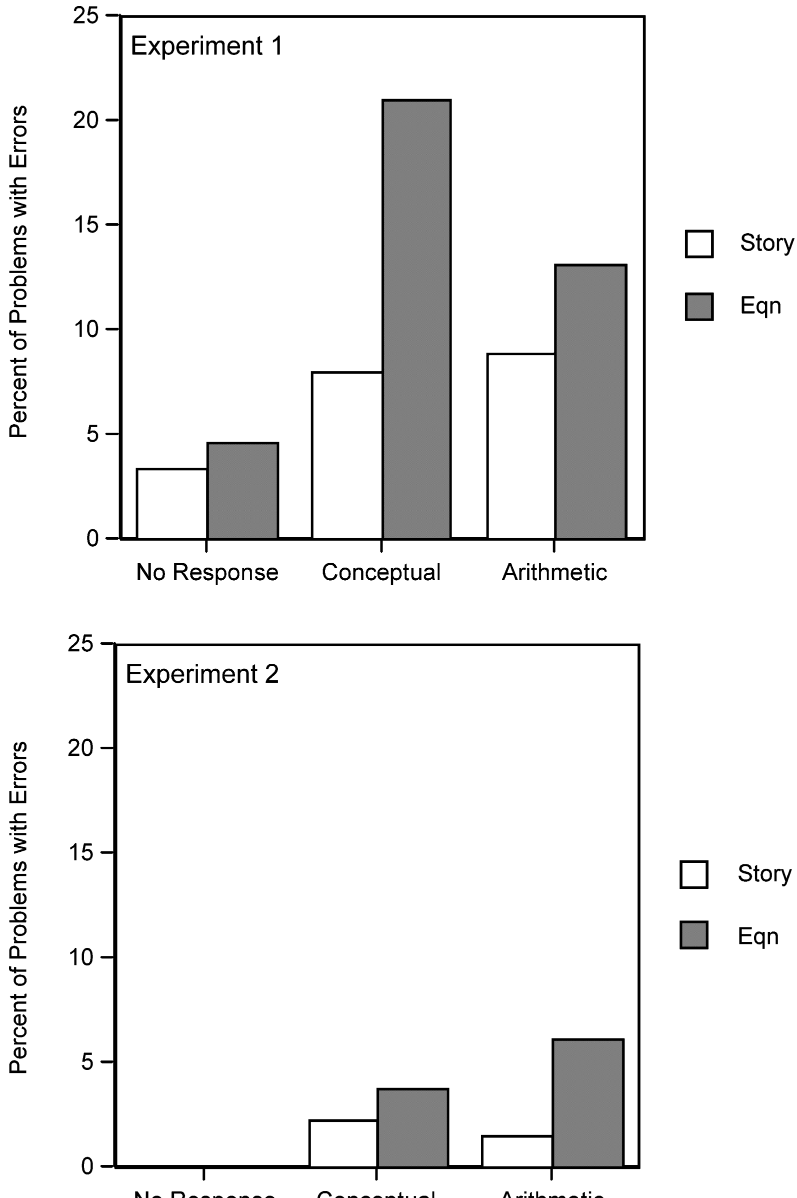


Fig. 6. Percentage of single-reference problems with errors of different types. In Experiment 1 (top), the verbal advantage is driven primarily by fewer Conceptual Errors on story problems than equations. In Experiment 2 (bottom), it is driven primarily by fewer Arithmetic Errors on story problems than equations.

As illustrated in Fig. 1, the particular complexity of double-reference problems makes the unwind strategy impossible to successfully apply. Although other informal strategies may work on double-reference problems, applying such strategies in the grounded story representation is more taxing on working memory than performing equation solving using abstract algebra symbols (we illustrate why below). Thus, we hypothesized that students would avoid informal strategies on double-reference story problems. As predicted, in Experiment 1, students used informal strategies on only 27% on the double-reference story problems but 55% of single-reference story problems. In Experiment 2, students used informal strategies on only 12% on the double-reference story problems but 62% of single-reference story problems.

When students do apply informal strategies to double-reference stories, they tend to make errors. The student solution in the first (crossed-out) attempt in Fig. 5d illustrates the typical outcome. On the discount problem, many students made inappropriate attempts to unwind by taking 15% of the sale price and adding the result to the sale price to get the original price. Such failures to represent the correct arithmetic operations were common when students attempted informal strategies on double-reference story problems (80% in Experiment 1, 25% in Experiment 2).

Fundamental to a theoretical explanation of the representation-complexity trade-off is this question: Why do students perform better on stories than equations for single-reference problems but worse for double-reference problems? Task analysis reveals that an informal solution is much more taxing on working memory when performed on a double-reference problem than on a single-reference problem.

Consider the number of mental transformations (cf. Mayer, 1982) needed to informally unwind a single-reference problem versus those needed to untangle a double-reference problem.[5] First consider the single-reference Lottery problem (see Fig. 1a). Focusing on the last part of the problem, the student needs to transform "some number divided by 3 gives 20.50" into "the number is 20.50 times 3." This transformation involves manipulating three quantities (some number, 3, and 20.5) and one operator, "divided by." In contrast, the mental transformations needed to informally untangle a double-reference problem like the Discount problem are much more numerous. The student needs to transform "some number minus .15 times the number" into "1 times some number minus .15 times the number" and eventually into ".85 times the number." This transformation involves three steps and, at the peak, manipulation of four quantities (1, .15, and two references to the number), and three operators (a "minus" and two "times"). Equations make double-reference problems easier because the analogous complex transformations need not be performed mentally but can be externalized in the concise equation representation: "$x - .15x = 1x - .15x = (1 - .15)x = .85x$."

As we saw, students appear sensitive to the differential costs and benefits of going abstract vs. staying grounded. They used formal algebra more often than not on the taxing double-reference problems but much less so on single-reference problems. Despite the extra difficulty of translating a story problem to an equation, students were more willing to pay this cost on double-reference than single-reference story problems.

7. General discussion

Our findings extend previous work that showed that, early in the development of algebra skill, students often succeed with grounded representations but fail with more abstract ones (Koedinger & Nathan, 2004). The present experiments have shown that, on more complex problems, students demonstrate an advantage for symbolic representations. Thus, we uncovered a representation-complexity trade-off, such that students show a verbal advantage on simple problems and a symbolic advantage on more complex problems. The present work generalizes past results by demonstrating a verbal advantage not only for younger elementary students (Baranes et al., 1989; Nunes et al., 1993) and beginning high school algebra students (Koedinger & Nathan, 2004) but also for students who are quite advanced in their mathematical skills (those in Experiment 2). More importantly, we identified boundary conditions on the verbal advantage by showing how increased problem complexity can yield a performance crossover whereby a symbolic advantage emerges.

Consistent with other recent efforts (Glenberg, Jaworski, Rischal, & Levin, 2007; Goldstone & Son, 2005; D. L. Schwartz, Martin, & Pfaffman, 2005; Sloutsky et al., 2005), we wished to move beyond the false dichotomy of whether grounded representations are generally better or worse than abstract representations. By decomposing the benefits along various cognitive dimensions and considering how these factors are affected by problem complexity and prior knowledge, we have taken steps toward an explanatory account of when, why, and how grounded and abstract representations provide benefits for performance. Our theoretical proposal for why the representation-complexity trade-off occurs (see Table 1) yields a set of hypotheses that are consistent with the results presented here. However, some alternative explanations for these results are possible, and we consider these below. We also discuss limitations of the current studies and open questions to be addressed in further research.

7.1. Benefits of grounded representations

One hypothesized benefit for grounded representations is that they put few demands on long-term memory because they involve familiar terms. But why should grounded representations be more familiar than abstract ones? The reason is because abstract formal representations tend to be used less frequently in everyday life, and they are usually introduced later in schooling. For instance, abstract representations for unknowns, like “x,” are used less frequently and encountered later in life than grounded representations for unknowns, like “how much.”

It is important to emphasize that the familiarity benefit of grounded representation does not require situated knowledge of particular story objects (Baranes et al., 1989; Nunes et al., 1993) and in particular predicts that familiar but less concrete representations will show benefits even for unfamiliar content. In addition to story problems, Koedinger and Nathan (2004) found that situation-less word problems (e.g., “Starting with some number, if I subtract 64 and then divide by 3, I get 20.5. What’s the number?”) were also solved more successfully than equations. Students had enough familiarity with the English descriptions of quantities and relations (e.g., “some number” is more familiar than “x”; “and then” is more familiar than the order of operations rules) that they did not exhibit the comprehension failures seen on equations.

The situation-less word problem manipulation also eliminates other hypotheses for why story problems are easier; for instance, that students have had more experience with story problems of particular types or overall forms (e.g., simple money problems) or that stories are more “embodied” or “enacted” (Glenberg et al., 2007; Lakoff & Núñez, 2000). The word problems we used are not of a single type and teachers report that they are not familiar in overall form (though they do involve familiar words and syntactic forms). Nor do these situation-less word problems involve embodiment and enactment, at least not as much as situated story problems.[6] Nevertheless, they are easier than equations.

Another test of the familiarity benefit would be to specifically manipulate familiarity within story problems. Cummins et al. (1988) found that particular words, like "altogether," cause elementary students greater difficulty in solving story problems than matched equations. Even for older students, with much better English knowledge, we have sometimes found particular story problems to be harder than others because of unfamiliar words, like "withhold."

While the familiarity benefit of grounded story representations was well addressed by Koedinger and Nathan (2004), the current studies highlight a second benefit. Grounded representations also help students to avoid and detect errors. The redundant semantic cues provided by grounded representations can support students in more reliable inferences about the problem structure and more reliable arithmetic performance.

We predicted that as a consequence of this reliability benefit, single-reference story problems would be easier than matched equations, even for students who have extensive algebra experience and thus are not lacking in familiarity with algebraic symbolism. This prediction was confirmed in Experiment 2 with highly skilled participants. Our error analysis further highlights how the two benefits play out at different stages in problem solving and acquisition. For low-skill students, reduction in No Response errors, due to the familiarity benefit, mostly accounts for better story performance (Koedinger & Nathan, 2004), whereas for high-skill students No Response errors disappear (due to high familiarity with algebra), but now a reduction in arithmetic errors, due to the reliability benefit, still yields better story performance (Fig. 6).

7.2. Is it just about translation?

Our theoretical analysis holds that performance on double-reference equations outstrips performance on double-reference story problems because equations are easier to manipulate and easier to externalize. More specifically, the conciseness of equations makes recombining unknown quantity relationships less taxing on working memory.

One might argue, instead, that equations provide benefit simply because they save participants from having to translate the word problems into equations themselves. Translation is indeed challenging. However, this line of reasoning is based on the false premise that story problem solving boils down to translating to an equation and solving it and, thus, story problems are necessarily harder than matched equations. As shown in Experiments 1 and 2, some story problems are actually easier than matched equations, and students use informal strategies as an alternative to the formal translate-and-solve strategy. Other kinds of story problems, such as double-reference problems, are harder than matched equations.

We have shown that double-reference equations are easier than double-reference stories. However, we have not directly shown that using equations makes double-reference story problems easier than when informal strategies are used. We could test this by requiring participants to solve story problems either with equations or informally and comparing performance under these conditions. We predict that on double-reference story problems, students will perform better when asked to solve using equations, whereas on single-reference story problems they will perform better when asked to solve informally. In support of this prediction, in the present study, when participants chose equation-solving on double-reference stories, they were much more successful than when they chose informal strategies (37 vs. 7% in Experiment 1; 80 vs. 55% in Experiment 2). We made efforts to address the potential selection bias inherent in this comparison by narrowing the analysis to students who used both strategies on different double-reference stories—these students showed the same benefit for equation solving. Nevertheless, an experimental follow-up would be worthwhile.

7.3. *Does the representation-complexity trade-off extend to other forms of complexity?*

In testing the representation-complexity trade-off hypothesis, we investigated a single source of complexity, namely, the unknown being referenced twice (double reference). There is evidence that the trade-off extends to other forms of complexity but not to all forms of complexity. The particular informal strategies that are afforded by a given grounded representation should continue to work unless there is a particular complex task demand that thwarts their use. For example, multiple references to the unknown thwarts unwind, and a requirement for precise solutions may thwart guess-and-test and graphical methods.

But not all forms of complexity thwart informal solutions and lead to crossover. Indeed, some single-reference problems are more complex on other dimensions than some double-reference problems. And in fact, the single-reference equations "$600 - 20 * x = 260$" and "$(X - 64) \div 3 = 20.50$" were solved less successfully than the double-reference equation "$x + (x + 6) = 38$." The forms of complexity in the single-reference problems (e.g., harder numbers, more kinds of arithmetic operators) are not the kind that thwart informal strategy use and demand abstraction.

Nevertheless, double reference is not the only form of complexity that we have found to bring out benefits of abstract representations. Verzoni and Koedinger (1997) found that for negative number problems, moving from one to two operations brought out the benefits of abstract representations. Stories were easier than equations for one-operator negative number problems, but the reverse held true for two-operator negative number problems. Combining negative numbers and two operators resulted in a symbolic advantage.[7] Similarly, Nathan and colleagues found that moving from linear problems to nonlinear problems (involving exponentials) also brought out benefits for abstract representations (Nathan, Stephens, Masarik, Alibali, & Koedinger, 2002).

We want to emphasize that the effect of complexity is not simply about preventing informal strategies but rather about the extra mental effort it takes or would take to address complex task demands within a grounded representation. As we illustrated, it is indeed possible to apply informal strategies on double-reference problems, but it is much more taxing on working memory to perform manipulations in the cumbersome, grounded verbal representation than it is in the concise, abstract equation representation.

In general, whether a particular form of complexity will cause a trade-off depends on the relationship between the nature of the complexity and the particular features of the strategies afforded by grounded and abstract representations. Table 1 suggests some relevant features, but we suspect this table will need to be extended to account for other benefits of grounded and abstract representations that might be revealed by other studies. Another row might be added, for instance, to characterize the purported benefit of abstraction that comes from uniformity, which lifts away from details of particular situations that can sometimes be misleading. Both Goldstone and Son (2005) and Sloutsky et al. (2005) demonstrated benefits of abstraction (or idealization) that are perhaps better explained by uniformity than externalizability or conciseness.

To illustrate in the algebra domain, one may be tempted by the grounded problem "If my investment increases by 10% one month and then decreases by 10% the next, where I am left?" to incorrectly answer "where you started." However, the abstract representation "Simplify $(X + .1X) - .1(X + .1X)$" is unlikely to yield the analogous error "X." The uniform or conventional mathematical operators lift away from the tempting particular associations of the everyday terms *increase* and *decrease*. Along these lines, in a closer item analysis of the implicit vs. explicit operator distinction on the discount problem in Experiment 1 (Table 3), we found that no student succeeded on the more grounded implicit operator problem, many reasoning informally and making the error of treating a decrease by 15% as the inverse of an increase by 15%. However, on the more abstract explicit operator version (where more abstract terms like *multiplied* and *subtracted* are used in addition to the grounded term *discount*), many students succeeded (14%) and they used formal equation solving to do so. The uniform mathematical operators prevent the incorrect inferences afforded by the terms used in the implicit-operator version of the problem. So, here we see a uniformity benefit of abstraction above and beyond the conciseness benefit that aids double-reference processing.

7.4. A developmental model of algebra problem-solving skill acquisition

Building on the present results, we propose a developmental model that specifies the sequence in which skills for solving problems are acquired in the domain of algebra. This sequence is (a) verbal single-reference problems, (b) symbolic single-reference problems, (c) symbolic double-reference problems, and finally (d) verbal double-reference problems. Early in the course of learning algebra, students tend to acquire skills for solving simple verbal problems before they acquire skills for solving comparable symbolic problems. Later in the course of learning algebra, students tend to acquire skills for solving complex symbolic problems before they acquire skills for solving comparable verbal problems.

Our proposed developmental model is reminiscent of others' research on shifts from grounded to abstract reasoning. For example, D. L. Schwartz and Black (1996) have demonstrated that, in learning to correctly solve problems about gears, people initially work with a grounded, depictive model (revealed in their spontaneous gesture production), and they eventually shift to a more abstract, computational strategy. Similarly, Case and colleagues have argued for a developmental progression from concrete, analog representations to more abstract, digital representations and integration of these representations (e.g., Kalchman, Moss, & Case, 2000). Indeed, the notion that concrete, grounded reasoning precedes and forms the basis for more abstract reasoning is one of the central tenets of Piaget's theory of intellectual development. However, although many studies have explored the advantages of grounded representations, until recently little research has focused on the complementary advantages of abstract representations (Goldstone & Son, 2005; D. L. Schwartz et al., 2005; Sloutsky et al., 2005).

7.4.1. Instructional implications

Developmental models like the one we have proposed can be useful in designing instruction, for instance, by providing guidance about the appropriate sequencing of topics in a curriculum. Such models can help curriculum designers to overcome the "expert blind spot" that occurs when experts' ideas about what may be difficult for students differ from reality (Koedinger & Nathan, 2004). Our model suggests that, at least for the early parts of learning algebra, the sequence of topics in most current algebra textbooks is inappropriate (Nathan, Long, & Alibali, 2002). Instead of learning to solve equations first and story problems later, our analysis suggests that it might be more effective for students to learn to solve simple story problems first and then bridge from this prior knowledge (Bransford, Brown, & Cocking, 1999) to a better understanding of equations.

Several experiments support this "bridging" idea (e.g., Brenner et al., 1997; Kalchman et al., 2000; Nathan, 1998; Nathan et al., 1992). For example, Koedinger and Anderson (1998) compared two versions of a computer-based cognitive tutor that contrasted a bridging approach with a textbook instructional approach to algebraic symbolization. In the bridging version, the cognitive tutor encouraged students to solve story problems numerically prior to attempting to translate the same problem to algebraic symbols. The cognitive tutor supports students in forming algebraic sentences (e.g., "42 * h + 35") by helping them to induce or generalize from examples of analogous arithmetic procedures (e.g., "42 * 3 + 35" and "42 * 4.5 + 35") they have just performed. In this situation, grounding is provided both by the situational context of the story problems and by working with concrete, numerical instances before moving to the abstract variable. The experimental comparison demonstrated that students learned more from this approach than from a textbook approach in which students symbolize prior to problem solving.

Understanding students' representational competencies and the order in which they are acquired is critical to effective instructional design. However, the relationship between grounded and abstract representations is not simply one of a progressive abstraction from grounded to abstract representations (cf. Bransford et al., 1999). This relationship is also important during more immediate problem solving, learning, and instructional events where a student may need to shift back and forth between grounded and abstract representations to take advantage of their complementary strengths (e.g., D. L. Schwartz & Black, 1996). Thus, from an instructional perspective, this relationship may be important both in the macrostructure of curriculum design and in the microstructure of lesson design.

Notes

1. How grounded a particular problem is for a particular student depends on that student's past experiences with the particular objects and events in a problem as well as the words or symbols used to refer to them. To the extent that there is substantial overlap in common experience across students, we can expect general effects of grounding. Our results suggest that this assumption holds.

2. We do not suggest that algebra students are more familiar with all ways of expressing quantitative relationships in English. Learning the mathematical meanings of some words and phrases, such as "profit" or "percent off," is indeed still difficult.
3. Although the unwind strategy does not work on double-reference problems, other informal strategies can be successful on such problems. For instance, by realizing that the sale price is 85% of the original, the student can, in effect, mentally combine the like terms x and 0.15x (this is a verbal reasoning analogy of the algebraic transformation of $x - 0.15x$ into .85x). This informal constraint combination, called *untangle* (see Table 3), is more sophisticated and memory demanding than the simpler unwind strategy.
4. Recall that the single-reference problems in Experiment 2 included both result-unknown and start-unknown problems, whereas in Experiment 1, all single-reference problems were start-unknowns. Students often used informal strategies, primarily arithmetic, to solve result-unknown stories (44 of 65, 68%). However, they used informal strategies, that is, not writing any new equations and going directly to the arithmetic, even more often on the corresponding equations (59 of 65, 91%).
5. Recall that our strategy analysis differentiates the informal unwind and untangle strategies (see Table 3). Untangle uniquely involves the combining of constraints needed to solve double-reference problems, analogous to combining like terms in an equation.
6. There is some level of enactment in terms like "starting with," "and then," and "I get"; however, in an unpublished study (Alibali, Koedinger, & Rose, in preparation) we found no performance decrement when these enactment terms were eliminated (e.g., "Some number minus 64 divided by 3 is 20.5. What's the number?").
7. This symbolic advantage for two-operator negative number problems differentiates from Koedinger and Nathan (2004) where two-operator positive number problems still show a verbal advantage.

Acknowledgments

The James S. McDonnell Foundation funded this research. Additional support was provided by the Undergraduate Research Initiative at Carnegie Mellon University. We are grateful to Mark Clark and Amy Foster for assistance with data collection and Michelle Rose, Nicole McNeil, and Erica Sueker for assistance in coding and establishing reliability.

References

Alibali, M. W., Koedinger, K. R., & Rose, M. (in preparation). What kinds of words lead to the verbal advantage in mathematical problem solving? Manuscript in preparation.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

Baranes, R., Perry, M., & Stigler, J. W. (1989). Activation of real-world knowledge in the solution of word problems. *Cognition and Instruction, 6*, 287–318.

Bednarz, N., & Janvier, B. (1996). Emergence and development of algebra as a problem-solving tool: Continuities and discontinuities with arithmetic. In N. Bednarz, C. Kieran, & L. Lee (Eds.), *Approaches to algebra: Perspectives for research and teaching*. (pp. 115–145). Dordrecht, The Netherlands: Kluwer Academic.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience and school*. Washington, DC: National Academy Press.

Brenner, M. E., Mayer, R. E., Moseley, B., Brar, T., Duran, R., Reed, B. S., et al. (1997). Learning by understanding: The role of multiple representations in learning algebra. *American Educational Research Journal, 34*, 663–689.

Carraher, T. N., Carraher, D. W., & Schliemann, A. D. (1987). Written and oral mathematics. *Journal for Research in Mathematics Education, 18*, 83–97.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 334–359.

Clement, J. (1982). Algebra word problem solutions: Thought processes underlying a common misconception. *Journal for Research in Mathematics Education, 13*, 16–30.

Collins, A., & Ferguson, W. (1993). Epistemic forms and epistemic games: Structures and strategies to guide inquiry. *Educational Psychologist, 28*, 25–42.

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology, 20*, 405–438.

Day, R. S. (1988). Alternative representations. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 261–305). San Diego, CA: Academic Press.

diSessa, A. A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior, 10*, 117–160.

Ernest, P. (1987). A model of the cognitive meaning of mathematical expressions. *British Journal of Educational Psychology, 57*, 343–370.

Geary, D. C. (1994). *Children's mathematical development: Research and practical applications*. Washington, DC: American Psychological Association.

Glenberg, A. M., Jaworski, B., Rischal, M., & Levin, J. R. (2007). What brains are for: Action, meaning, and reading comprehension. In D. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 221–240). Mahwah, NJ: Lawrence Erlbaum.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences, 14*, 69–110.

Greeno, J. G., & Hall, R. P. (1997, January). Practicing representation: Learning with and about representational forms. *Phi Delta Kappan, 78*, 361–367.

Hall, R. P., Kibler, D., Wenger, E., & Truxaw, C. (1989). Exploring the episodic structure of algebra story problem solving. *Cognition and Instruction, 6*, 223–283.

Heffernan, N., & Koedinger, K. R. (1997). The composition effect in symbolizing: The role of symbol production versus text comprehension. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (pp. 307–312). Mahwah, NJ: Lawrence Erlbaum.

Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. *Child Development, 54*, 84–90.

Kalchman, M., Moss, J., & Case, R. (2000). Psychological models for the development of mathematical understanding: Rational numbers and functions. In S. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 1–38). Mahwah, NJ: Lawrence Erlbaum.

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

Kirshner, D. (1989). The visual syntax of algebra. *Journal for Research in Mathematics Education, 20*, 274–287.

Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. C. (2005). Middle school students' understanding of core algebraic concepts: Equality and variable. *Zentralblatt für Didaktik der Mathematik – International Reviews on Mathematical Education, 37*, 68–76.

Koedinger, K. R., & Anderson, J. R. (1998). Illustrating principled design: The early evolution of a cognitive tutor for algebra symbolization. *Interactive Learning Environments, 5*, 161–180.

Koedinger, K. R., & MacLaren, B. A. (2002). *Developing a pedagogical domain theory of early algebra problem solving* (CMU-HCII Tech. Rep. 02-100). Pittsburgh, PA: Carnegie Mellon University.

Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representation on quantitative reasoning. *Journal of the Learning Sciences, 13*, 129–164.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology, 17*, 248–294.

Küchemann, D. E. (1978). Children's understanding of numerical variables. *Maths in Schools, 7*, 23–26.

Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.

Lewis, C. H. (1981). Skill in algebra. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 85–110). Hillsdale, NJ: Lawrence Erlbaum.

Matz, M. (1980). Toward a computational theory of algebra competence. *Journal of Mathematical Behavior, 3*, 93–166.

Mayer, R. E. (1982). Different problem-solving strategies for algebra word and equation problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 448–462.

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28*, 451–466.

Nathan, M. J. (1998). Knowledge and situational feedback in a learning environment for algebra story problem solving. *Interactive Learning Environments, 5*, 135–159.

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra word problem comprehension and its implications for the design of computer learning environments. *Cognition and Instruction, 9*, 329–389.

Nathan, M. J., Long, S. D., & Alibali, M. W. (2002). Symbol precedence in mathematics textbooks: A corpus analysis. *Discourse Processes, 33*, 1–21.

Nathan, M. J., Stephens, A. C., Masarik, D. K., Alibali, M. W., & Koedinger, K. R. (2002). Representational fluency in middle school: A classroom study. In D. S. Mewborn, P. Sztajn, D. Y. White, H. G. Wiegel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of the Twenty-Fourth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 462–472). Columbus, OH: ERIC Clearinghouse for Science, Mathematics and Environmental Education.

Nhouyvanisvong, A. (1999). *Enhancing mathematical competence and understanding: Using open-ended problems and informal strategies*. Unpublished doctoral dissertation. Pittsburgh, PA: Carnegie Mellon University.

Nunes, T., Schliemann, A. D., & Carraher, D. W. (1993). *Street mathematics and school mathematics*. New York: Cambridge University Press.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Paivio, A., Clark, J. M., & Khan, M. (1988). Effects of concreteness and semantic relatedness on composite imagery ratings and cued recall. *Memory & Cognition, 16*, 422–430.

Palmer, S. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Lawrence Erlbaum.

Payne, S. J., & Squibb, H. R. (1990). Algebra mal-rules and cognitive accounts of error. *Cognitive Science, 14*, 445–491.

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.

Schwartz, D. L., & Black, J. B. (1996). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science, 20*, 457–497.

Schwartz, D. L., Martin, T., & Pfaffman, J. (2005). How mathematics propels the development of physical knowledge. *Journal of Cognition and Development, 6*(1), 65–86.

Schwartz, S. H. (1971). Modes of representation and problem solving: Well evolved is half solved. *Journal of Experimental Psychology, 91*, 347–350.

Schwartz, S. H. (1972). Representation in deductive problem solving: The matrix. *Journal of Experimental Psychology, 95*, 343–348.

Scribner, S. (1986). Thinking in action: Some characteristics of practical thought. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence* (pp. 13–30). Cambridge, UK: Cambridge University Press.

Sleeman, D. (1986). Introductory algebra: A case study of student misconceptions. *Journal of Mathematical Behavior, 5*, 25–52.

Sloutsky, V. M., Kaminski, J. A. & Heckler, A. F. (2005). The advantage of simple symbols for learning and transfer. *Psychonomic Bulletin & Review, 12*, 508–513.

Verzoni, K., & Koedinger, K. R. (1997). *Student learning of negative number: A classroom study and difficulty factors assessment*. Paper presented at the annual meeting of the American Educational Research Association, April, Chicago, IL.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science, 21*, 179–217.