COGNITION AND INSTRUCTION, 17(3), 215–248

Multimedia-Supported Metaphors for Meaning Making in Mathematics

Roxana Moreno and Richard E. Mayer
Department of Psychology
University of California, Santa Barbara

We examined a number-line metaphor presented using interactive multimedia as a means of helping students build connections between an arithmetic procedure and their existing conceptual knowledge. Elementary school children learned to add and subtract signed numbers through a computer-based multimedia program over 4 training sessions. Participants received 64 example problems presented only in symbolic form (single-representation [SR] group) or in symbolic, visual, and verbal forms (multiple-representation [MR] group). In Experiment 1, compared to the SR group, the MR group (a) showed a larger pretest-to-posttest gain for high-achieving students but not for low-achieving students, (b) showed a greater gain on difficult problems but not easy problems, (c) learned faster during training, and (d) showed a greater pretest-to-posttest reduction in the use of conceptual bugs reflecting conceptually confusion between negative signs and subtraction operators. In Experiment 2, high spatial ability students in the MR group outperformed low spatial ability students on pretest-to-posttest gain. Productive learning with MRs is strongest when working memory is not overloaded, so cognitive load and MR theories can be reconciled.

What does it mean to understand a computational procedure in mathematics? For example, what can be done to help students make sense out of procedures for addition and subtraction of signed whole numbers, such as $2 - -5 = \_\_\_$? In this study we examine the classic instructional proposal (Brownell, 1928, 1935; Resnick & Ford, 1981) that students understand arithmetic procedures by mentally linking them with appropriate concrete models.

There is a growing consensus among cognitive scientists and mathematics educators that learning of computational procedures is an active process in which stu-

Requests for reprints should be sent to Roxana Moreno or Richard E. Mayer, Department of Psychology, University of California, Santa Barbara, CA 93106. E-mail: moreno@psych.ucsb.edu or mayer@psych.ucsb.edu

dents seek to make sense out of the to-be-learned procedures by constructing mental models (English, 1997; Grouws, 1992). In this study we focus on the number line as an appropriate model or grounding metaphor for signed arithmetic because it has been implicated as a central conceptual structure underlying mathematical cognition (Case & Okamoto, 1996). Based on recent advances in multimedia learning (Mayer, 1997), we examine the effects of depicting the execution of computational procedures as computer-supported visual and verbal representations of movements along a number line.

In the remainder of this section, we examine the questions of why to use metaphors to foster learning of mathematical procedures, which metaphors are the best ones to use, and how to use metaphors within a computer-supported multimedia environment to foster meaningful learning. We then provide a cognitive theory of multimedia learning to guide our understanding of how people process multiple representations (MRs), and we derive predictions based on the theory.

WHY USE METAPHORS TO TEACH ABOUT ARITHMETIC PROCEDURES?

Two Views of Metaphors

Metaphors may be viewed as unnecessary adjuncts to serious mathematics learning or as essential instructional tools for fostering mathematical understanding. For example, a recent review of mathematics textbooks revealed two predominant types of instruction for signed computation: (a) abstract and disembodied approaches that emphasize symbol manipulation based on seemingly arbitrary rules such as, "If there are two minus signs, then add the numbers," and (b) concrete and embodied approaches that metaphorically relate computation to concrete models such as walking along a path (Mayer, Sims, & Tajika, 1995). Interestingly, the first approach, focusing on presenting problems mainly in symbolic form, dominated U.S. textbooks, whereas the second approach, relating the symbolic representation to graphic- and text-based models, was rare in U.S. textbooks (Mayer et al., 1995).

Why have metaphors been ignored in traditional mathematics education? Fuson (1992a) showed how traditional mathematics instruction is based on the conception of "teaching children to be doers rather than thinkers" and thereby "considerably underestimates what children can learn" (p. 57). Similarly, Sowder (1992) observed that "students are not disposed to make sense out of numbers" because "they do not see school mathematics as a sense-making activity" but rather as a "collection of facts and rules to be memorized" (p. 20). Bassok (1997) also showed how school children "learn that, while in school, they will fare better if they completely ignore their world knowledge" (p. 222) so they learn to apply computational procedures blindly.

The decision to use metaphors in mathematics education may depend on one's conception of learning in general, and learning of computational procedures in particular. For example, the search for ways to help students learn computational procedures can be influenced by whether one views *learning as rule building* (Anderson, 1993; Newell & Simon, 1972), in which learners apply general rules to problems, or *learning as model building*, in which learners construct mental models of problem situations (English, 1997; Fuson 1992a, 1992b; Goswami, 1992; Halford, 1993; Hiebert & Carpenter, 1992; Kintsch & Greeno, 1985). The rule-building view has a long and influential history in both psychology and education, dating back to the classic work of Thorndike (1913) on learning cognitive skills by drill and practice and reflected in the modern work of Anderson (1993) on representing mathematical skills as general production rules.

In contrast, the model-building view can be seen in classic Gestalt theories of mathematical problem solving (Wertheimer, 1959), in Piaget's (1952) theories of the development of mathematical competence, and in Brownell's (1928, 1935) meaning theory of arithmetic, which emphasized the role of concrete manipulatives (Resnick & Ford, 1981) and is reflected in modern research on the situational context of mathematical cognition (Nunes, Schliemann, & Carraher, 1993). According to the model-building view, learning is a constructive process of sense making in which learners make connections between new information and existing conceptual models. The model-building approach is based on a "move away from the traditional notion of reasoning as abstract and disembodied to the contemporary view of reasoning as embodied and imaginative" (English, 1997, p. vii).

Metaphor as an Aid to Mathematical Understanding

For the last 25 years, scholars have repeatedly argued for the value of metaphors as aids to cognition (Gentner & Stevens, 1983; Lakoff & Johnson, 1980; Ortony, 1993; Vosniadou & Ortony, 1989), with special emphasis on the role of metaphors as aids to scientific and mathematical understanding (English, 1997; Glynn, Duit, & Thiele, 1995). English (1997) argued that metaphor—especially when combined with imagery—can support students' construction of mental models that are fundamental to mathematical reasoning but noted that "insufficient attention has been given to the important role these play in mathematical reasoning" (p. 4). English posited that "image-based reasoning in mathematics ... has not received the attention it deserves largely because mathematics traditionally has been viewed as a purely abstract discipline" (p. 10).

The strongest form of the learning-as-model-building view is reflected in Lakoff and Johnson's (1980) contention that "our ordinary conceptual system ... is fundamentally metaphorical in nature" (p. 3), and English's (1997) proposal that "reasoning with metaphors is considered a fundamental way of human thinking and communication" (p. 7). Accordingly, the model-building view is particularly relevant for mathematics learning: "Metaphor is central to the structure of mathematics and to our reasoning with mathematical ideas," and "metaphorical reasoning can assist students in their interpretations of formal mathematical language" (English, 1997, p. 7). Alexander, White, and Daugherty (1997) stated that "analogical reasoning is foundational to learning, in general, and to early mathematics learning in particular" (p. 117).

In summary, the rationale for using metaphors in mathematics instruction is that metaphors can be aids to mathematical understanding. In this article, we adopt a model-building view and seek to understand how multimedia environments can introduce metaphors that foster the construction of appropriate mental models in young learners.

WHICH METAPHOR TO USE FOR TEACHING ABOUT ARITHMETIC COMPUTATION?

It is one thing to acknowledge the potential of metaphor as an aid to mathematical understanding, and it is quite another to identify the appropriate metaphor to help youngsters understand computational procedures for addition and subtraction of signed numbers. Fortunately, several converging research programs suggest that the most likely candidate for an appropriate metaphor is the *number line*. A number line consists of a line with numbers as discrete points along the line; in its most complete form 0 is at the center with consecutive positive integers to the right and decreasing as consecutive negative integers to the left. The number-line metaphor is implicated in developmental research on central conceptual structures underlying number sense (Case & Okamoto, 1996), linguistic analyses of grounding metaphors in arithmetic (Lakoff & Nunez, 1997), and educational research on concrete manipulatives for understanding arithmetic (Hiebert & Carpenter, 1992).

Number Line as a Central Conceptual Structure

Case and his colleagues (Case & Okamoto, 1996; Griffin, Case, & Capodilupo, 1995) have shown that learning arithmetic procedures such as addition and subtraction of signed numbers must be tied to the development of children's central conceptual structures—that is, cognitive representations that the child uses to understand new situations. According to Case and Okamoto (1996), the mental number line is the central conceptual structure underlying children's learning of arithmetic. Importantly, students who enter elementary school without the concept of a number line tend to have difficulty in learning arithmetic: "A surprising proportion of children from low-income ... families ... do not arrive in school with central cognitive structures in place" so "their first learning of addition and subtraction may be a meaningless experience" (Griffin & Case, 1996, p. 102).

Case and his colleagues (Griffin & Case, 1996; Griffin et al., 1995) described instructional activities aimed at fostering children's development of the concept of a number line, including playing board games involving moving a token along a number-line path. Students who received number-line training showed an improvement in their knowledge of a mental number line, and were more successful in subsequently learning arithmetic than students who did not receive the training. In reviewing the beneficial effects of helping students develop a number-line concept, Bruer (1993) noted that students who lack the number-line concept must "try to understand school math as a set of arbitrary procedures" (p. 90). Bruer further noted that "for mathematics to be meaningful, conceptual knowledge and procedural skills have to be interrelated in instruction" (p. 90). According to this line of research, the corresponding conceptual knowledge for arithmetic procedures is the mental number line.

Number Line as a Grounding Metaphor

In their linguistic analysis of metaphors that promote mathematical cognition, Lakoff and Nunez (1997) showed that a grounding metaphor helps a student to relate mathematical ideas to everyday experience, and a fundamental grounding metaphor for understanding arithmetic procedures is "arithmetic is motion" (p. 37). According to the arithmetic-is-motion metaphor, "numbers are locations on a path," "the mathematical agent is a traveler along that path," "arithmetic operations are acts of moving along the path," and "the result of an arithmetic operation is a location on the path" (p. 37). Furthermore, according to the arithmetic-is-motion metaphor "addition of a given quantity is taking steps ... forward," whereas "subtraction of a given quantity is taking steps ... backward" (p. 37). The evidence for this assertion is derived from linguistic examples that entail the idea of moving along a number line: "How *close* are these two numbers? 37 is *far away from* 189712. The result is *around* 40. Count up to 20, without *skipping* any numbers" (p. 37).

Lakoff and Nunez's (1997) analysis reveals that all grounding metaphors are not equally useful for understanding addition and subtraction of signed numbers. For example, a commonly used metaphor to explain signed arithmetic is the idea that a negative number is made of anti-matter (Lakoff & Nunez, 1997; Mayer et al., 1995), such that $3 + -3 = 0$ means that the -3 annihilates the 3. Lakoff and Nunez (1997) concluded that the anti-matter metaphor is not a good grounding metaphor for signed arithmetic but the number line is:

> The easiest natural extension of one of the grounding metaphors to negative numbers is the motion metaphor ... addition and subtraction of negative numbers can then be

given a relatively easy extension of the metaphor: when you encounter a negative number, turn around in place. (p. 39)

In searching for a metaphor for arithmetic, English (1997) suggested "a line metaphor to represent our number system" in which "numbers are considered as points on a line" (p. 8). English proposed that practice in coordinating visual and symbolic representations is needed to help students ground arithmetic procedures within the number-line metaphor: "Although it can be an effective metaphor for our number system, the number line is a complex representation requiring students to integrate two forms of information, namely, visual and symbolic" (p. 8). Although widely recognized as the best available grounding metaphor for addition and subtraction of signed numbers, the number-line metaphor has limits. For example, if a student views arithmetic as moving along a series of discrete stepping stones, he or she may have difficulty learning to add and subtract fractions.

Number Line as a Concrete Manipulative

Following earlier work by Brownell (1928, 1935) on the role of concrete manipulatives, Hiebert and Carpenter (1992) and Fuson (1992b) have shown how mental representations—including concrete manipulatives—are at the heart of mathematical cognition. However, for concrete manipulatives to be effective, they must make sense to students—a requirement that can be met through active discussion and experimentation. Resnick (1983) has reviewed research showing that "by the time they enter school most children have already constructed a representation of number that can be appropriately characterized as a mental number line" (p. 110–111). This informal knowledge can be used to accomplish a "considerable amount of arithmetic problem solving" (p. 111). It follows that a number line is a primary candidate for a concrete manipulative in teaching of arithmetic procedures.

For example, Lewis (1989) developed an instructional program to help college students represent arithmetic word problems in which students translated the problem into a graphic representation on a number line. Students learned to place letters on a number line, corresponding to variables in the problem. The training resulted in substantial improvements in mathematical problem solving as compared to students who did not receive instruction in how to use a number-line diagram. More recently, Brenner et al. (1997) developed a program to help middle-school students translate word problems and tables into graphs—a sort of two-dimensional number line. Students who received this experience showed greater improvements in their mathematical problem-solving skills than students who received conventional instruction.

In using a concrete manipulative, "the teacher is in effect creating a metaphor for the child to use as an assimilation paradigm" (Davis & Maher, 1997, p. 100); furthermore, these "representations—sometimes mental and sometimes on paper—make it possible to think about some idea which might otherwise be labeled as new" (p. 114).

HOW TO HELP STUDENTS BUILD LINKS BETWEEN CONCEPTUAL AND PROCEDURAL KNOWLEDGE?

In this research we focus on one aspect of meaningful mathematics learning: encouraging students to integrate new arithmetic procedures with their existing conceptual knowledge. To accomplish this goal, we rely on the role of metaphor in learning, and, in particular, we focus on the number line as a central conceptual structure, grounding metaphor, or assimilating paradigm. In this section, we describe one implementation of the number-line metaphor within a computer-supported multimedia environment. We are intrigued by the possibility that multimedia can offer a powerful environment for helping students build connections among MRs of arithmetic problems. In particular, we examine how multimedia environments can help students build connections between formal computational procedures using symbols—that is, rule-based symbol manipulation—and informal conceptual knowledge about moving along a path—that is, visual and verbal representations of the procedure. In choosing to focus on computer-supported learning, we recognized that the number-line metaphor can be used in a variety of educational venues, and we do not purport that there is anything magical about multimedia per se. Instead, we seek to understand the conditions under which students build and coordinate MRs in a multimedia environment.

How can we use multimedia environments to help students build connections between mathematical procedures and their existing conceptual knowledge? Multimedia environments can introduce students to MRs of the same concept or procedure (such as presenting text, graphics, animations, sound, and video) and allow students to manipulate and coordinate these MRs within computer microworlds (Mayer, 1997). Moreover, multimedia environments have the capability of creating dynamic representations of constructs that are frequently missing in the mental models of novices (Kozma, 1991). In this study, we examined the cognitive consequences of learning to solve 64 signed arithmetic problems presented in one form of representation (i.e., symbolic) or three coordinated forms of representation (i.e., symbolic, visual, and verbal) within an interactive multimedia environment.

Learning an Arithmetic Procedure With a Single Representation

Students may learn an arithmetic procedure, such as adding and subtracting signed integers, by working on example problems presented in a standard single form. Fig-

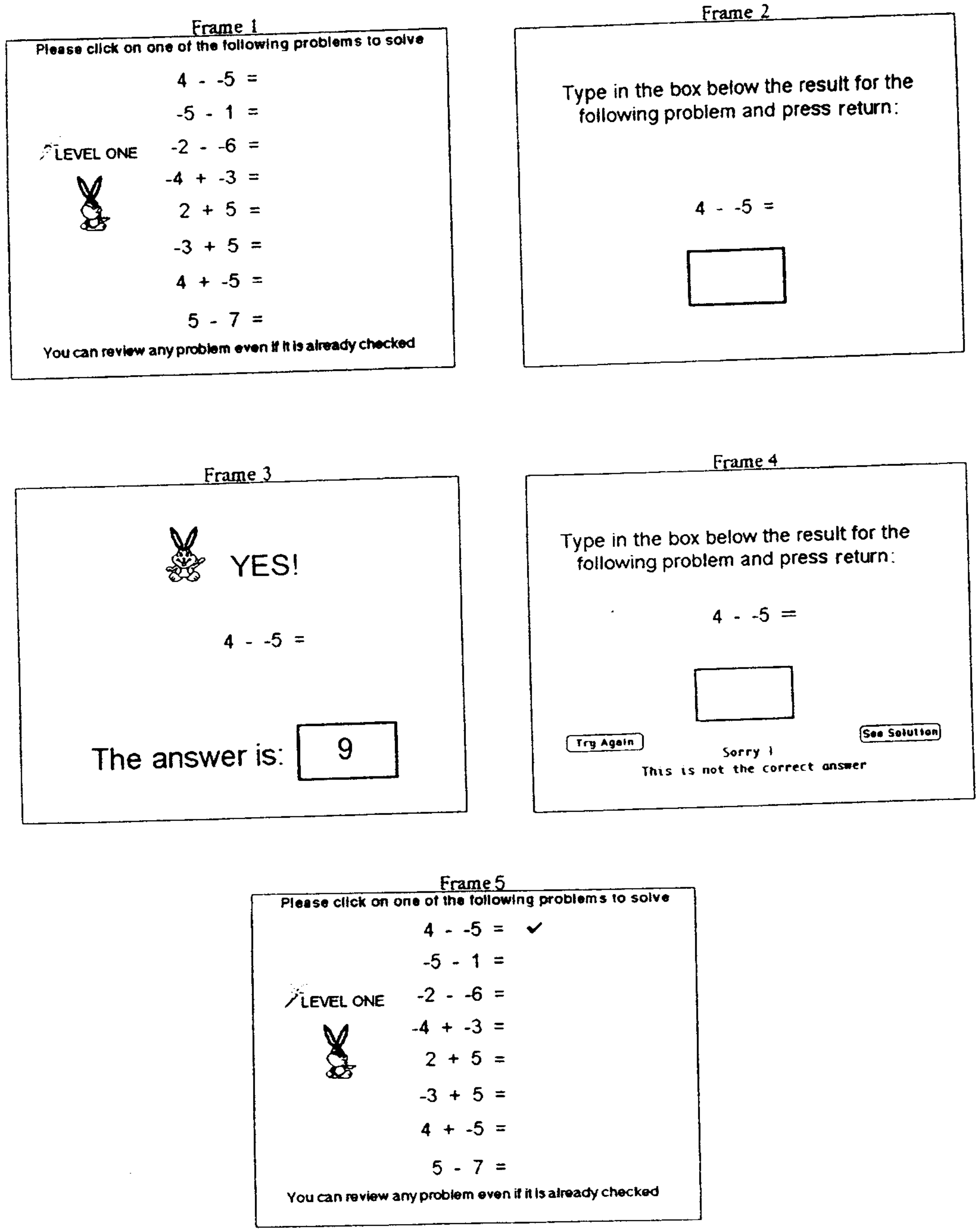

FIGURE 1 Selected frames from the single representation program.

ure 1 shows selected frames from an example problem presented only in symbolic form (single-representation [SR] group). In the first frame, the student selects one of eight problems to solve from a problem menu. In the second frame, a problem appears using only numerals, signs, and operator symbols, and the student is asked to type in a numerical answer using a keypad. In the third frame, if the student types in the correct numeral (and sign, if negative), a smiling bunny appears on the screen along with the a large "YES" and the completed number sentence. In the fourth frame, if the student types in the wrong number, the student is given the chance to try again (by typing on the keypad) or to see the correct number. If the student clicks on "Try Again," the second frame appears again; if the student clicks on "See Solution," the correct numerical answer appears in the box. In the fifth frame, after either producing or asking to see the correct number, the student returns to the problem menu, with a check now placed next to the completed sample problem. The instructional program designed for the SR group uses the most common format of arithmetic instruction in the United States—that is, the traditional, practice-based sessions in which students were only presented with the symbolic representation of a number sentence such as 2 – –3 = ___ and had no additional instruction or explanation (Mayer et al., 1995).

Learning an Arithmetic Procedure With MRs

In contrast to learning with an SR, students may learn by seeing and coordinating MRs. For example, Figure 2 shows selected frames from an example problem presented in symbolic, visual, and verbal form (MR group). First, the student selects one of eight problems to solve from the same problem menu as for the SR group (as shown in the first frame of Figure 1). Then, as shown in the first frame in Figure 2, the student sees the problem presented in symbolic form (as 4 – –5 =) and a box as in the SR group. In addition, however, the screen also contains a number line showing integers from –9 to 9 with a bunny standing at the 0 point and a simulated joystick consisting of four alternatives that make the bunny face to the left, face to the right, jump forward one step, or jump backward one step. The student may click on any combination of the four joystick options and instantly see the resulting change in the bunny on the number line; the student is instructed to try to figure out the problem by moving the bunny along the number line using the joystick. When the student is ready to answer, the student types in a numeral (and negative sign, if needed).

If the student's answer is correct, the bunny appears along with a large "YES" as in the SR group; in addition, this is followed by an animated sequence that consists of four major steps. First (in the second frame in Figure 2), the symbol "4" is highlighted, on-screen text appears stating "FIRST: FIND MY STARTING POINT 4 means GO TO 4," and then (in the third frame) the bunny

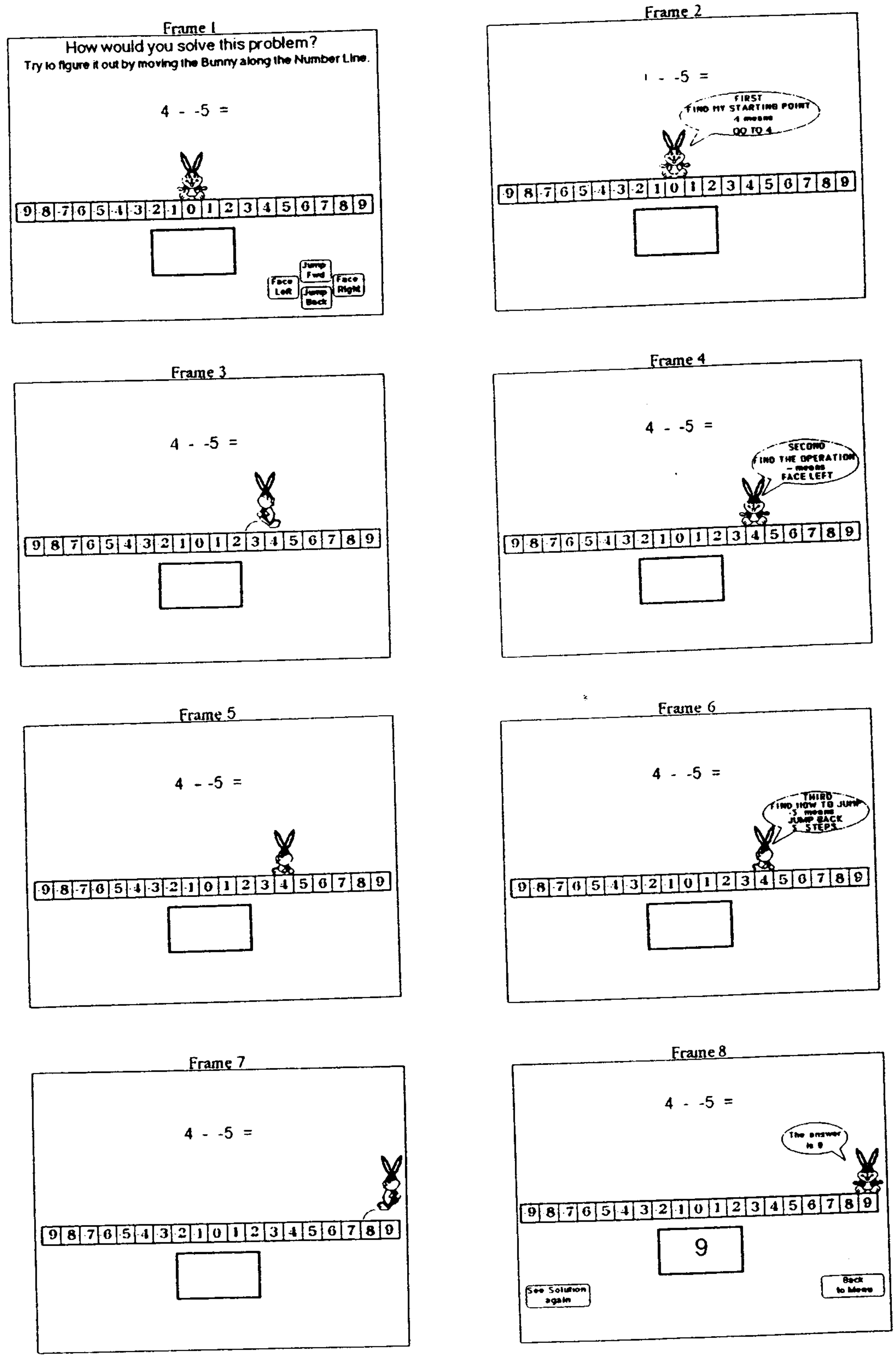

FIGURE 2 Selected frames from the multiple representations program.

moves to position 4 on the number line. Thus, the learner sees how the first step in the procedure can be represented in symbols, words, and pictures. Second (in the fourth frame), the minus sign is highlighted, on-screen text appears stating, "SECOND: FIND THE OPERATION – means FACE LEFT," and then (in the fifth frame) the bunny turns to face the left side of the screen. Third (in the sixth frame), the symbol "–5" is highlighted, on-screen text appears stating, "THIRD: FIND HOW TO JUMP –5 means JUMP BACK 5 STEPS," and then (in the seventh frame), the bunny makes five jumps to the right. Finally (in the eighth frame), the entire number sentence "4 – –5 = " is highlighted, the bunny stands on the 9 position, and on-screen text says "The answer is 9." At this point, the student may click on "See Solution Again," which repeats this four-step animated sequence, or on "Back to Menu," which takes the student back to the menu frame (as shown in the last frame of Figure 1).

If the student's answer is not correct, the student may click on "Try Again" (which returns the student to the first frame in Figure 2) or "See Solution" (which presents the same four-step animated sequence as depicted in frames 2 through 8 in Figure 2). After either producing or asking to see the correct answer, the student returns to the problem menu as in the SR group (fifth frame in Figure 1). The instructional program designed for the MR group is one possible instantiation of the mathematics reform literature's calls for infusing MRs based on the number-line metaphor in the teaching of arithmetic procedures. The contrast between SR and MR programs resides in the fact that the former provides students with symbols and no further explanations, whereas the latter provides students with symbols plus additional visual and verbal explanations intended to help students advance their understanding and learning.

TWO APPARENTLY OPPOSING THEORIES OF MULTIMEDIA LEARNING

Although research in learning with MRs in a multimedia environment is still evolving, recent studies in the area of instructional design have started to show the tension between two seemingly opposing theories: MR theory, which builds on Paivio's dual-coding theory (Clark & Paivio, 1991; Paivio, 1986) and cognitive-load theory, which builds on Baddeley's (1992) model of working memory. Although MR theory suggests that teaching with more representations facilitates and strengthens the learning process by providing several mutually referring sources of information (Kozma, Russell, Jones, & Marx, 1996; Mayer, 1997; Mayer & Anderson, 1991), cognitive-load theory suggests that the mental integration of the disparate sources of information may generate a heavy cognitive load that is detrimental to learning (Chandler & Sweller, 1991; Sweller, 1988, 1989; Sweller, Chandler, Tierney, & Cooper, 1990).

Cognitive load may be affected either by the intrinsic nature of the material to be learned (intrinsic cognitive load) or by the instructional design used to teach the material (extraneous cognitive load). Intrinsic cognitive load is dependent on the amount of interactivity of the material to be learned (Marcus, Cooper, & Sweller, 1996), which is a function of the number of elements that have to be processed simultaneously in working memory. On one hand, tasks low in interactivity have elements that can be learned independently and serially. On the other hand, tasks high in interactivity can not be understood unless several elements are held and processed simultaneously in working memory. Extraneous cognitive load is solely the result of a particular instructional design and occurs when the design imposes a load that bears little relation to the learning process. For example, Sweller (1988) found that to use means–ends analysis—which requires the learner to consider the initial problem state, the future goal state, and the difference between the two states—the operator needed to reduce the difference, and the resulting subgoals impose heavy demands on working memory. The alternative method of providing the students with worked examples has proved to be significantly more efficient in many studies (Cooper & Sweller, 1987; Paas, 1992; Trafton & Reiser, 1993; Zhu & Simon, 1987).

MR theory and cognitive-load theory seem to have opposing predictions and, therefore, different practical implications. Whereas MR theory encourages the development of multimedia programs that concurrently present the same concepts or procedures in several forms, cognitive-load theory encourages the development of multimedia programs that do not overload the learner with too much redundant concurrent information.

A goal of this study is to examine the apparent conflict between MR theory and cognitive-load theory. To accomplish this goal, we compared the cognitive consequences of learning how to add and subtract signed numbers in an interactive multimedia environment when example arithmetic problems were presented solely in symbolic form (SR group) or in symbolic, visual, and verbal forms (MR group). According to MR theory, students trained with example problems presented in three different forms encode the material more deeply than those trained with problems in only symbolic form. Thus, the MR group should show better pretest-to-posttest gains in solving signed arithmetic problems than the SR group. Cognitive-load theory predicts the opposite result. The high interactivity of the material involved in the learning process entails a high intrinsic cognitive load, so extra forms of representation will add to the total amount of cognitive load, resulting in less learning. According to cognitive-load theory, students in the SR group should outperform students in the MR group on pretest-to-posttest gain in solving signed arithmetic problems.

Our study adds an extraneous cognitive load by using MRs in a multimedia environment. We chose material that is very high in element interactivity, that is, it is not possible to understand how to add and subtract signed numbers by solely learning the elements involved in sequence. Only by simultaneously studying the specific interaction of each operand and operation can a student successfully build a model to solve all problem types, so the instructional material is high in intrinsic cognitive load. When the material to be learned is already high in intrinsic cognitive load, adding extraneous cognitive load may be detrimental to learning by exhausting the working memory capacity. Therefore, the task of addition and subtraction of integers provides the necessary learning conditions to test the implications of cognitive load theory.

PREDICTIONS

Predictions Based on MR Theory

Based on MR theory, students who learned with MRs should show a greater improvement in their ability to add and subtract signed numbers than students who learned with an SR. We measure improvements in three ways: pretest-to-posttest gains in the number of correct answers in solving 18 signed arithmetic problems, the pattern of improvement in the number of correct answers for 16 equivalent example problems across the four learning trials, and changes in the types of strategies (or bugs) used to generate incorrect answers on the pretest and posttest. Thus, MR theory predicts that compared to the SR group, the MR group will show a larger pretest-to-posttest gain, faster rate of learning across the four learning sessions, and more pretest-to-posttest transitions from conceptually poor to conceptually rich bugs for incorrect answers. These effects may be particularly strong for difficult problems because these offer the most room for improvement.

Predictions Based on Cognitive-Load Theory

A straightforward interpretation of cognitive-load theory generates the opposite set of predictions, that is, the SR group should improve more than the MR group on each of the dependent measures. If the amount of information presented in the MR treatment overloads working memory, learning may be hindered. If the SR treatment does not overload working memory and provides sufficient information for learning the procedure, it can produce better learning than the MR treatment on each of the three measures described previously.

Predictions Based on MR Theory Combined With Cognitive-Load Theory

Can the two competing theories be reconciled? Based on a combination of MR and cognitive-load theory, a third set of predictions is that the positive effects of MR learning should be strong for high-achieving students or high-spatial students but not for low-achieving or low-spatial students. The rationale for this set of predictions about arithmetic achievement is that higher achieving students have already automated their basic arithmetic skills and, therefore, have more cognitive resources to create the connections among the visual, verbal, and symbolic representations. In contrast, lower achieving students must devote some of their limited working memory resources to learn the basic arithmetic operations, leaving less resources for making connections among MRs. The rationale for this set of predictions about spatial ability is that high-spatial students do not need to use cognitive resources to hold images in working memory, so they may devote those limited resources to building connections among visual, verbal, and symbolic representations. In contrast, low-spatial ability learners must devote more cognitive resources to holding images of the bunny's movement in working memory, so they have less resources to use on building connections among MRs. Overall, the best strategy for low-achieving and low-spatial ability learners may be to focus only on the symbolic form of representation, a tactic that would render the SR and MR treatments equivalent.

EXPERIMENT 1

Our purpose in Experiment 1 was to examine the cognitive consequences of learning how to solve signed arithmetic problems in an interactive multimedia environment under two different methods of instruction. One group of students (SR group) studied 16 problems in each of four training sessions with the symbolic number sentence being the only representation available. Another group of students (MR group) studied the same problems in four training sessions but had (a) the same traditional symbolic representation used for the SR group, (b) a visual representation that uses a number line and a computer animation to show how the symbolic number sentence relates to a bunny's movements along the number line, and (c) a verbal representation that uses written explanatory text that has the bunny describe in words how the symbols relate to its movements along the number line.

METHOD

Participants and Design

The participants were 60 sixth-grade students from an elementary school in Southern California who lacked substantial prior knowledge about addition and subtraction of signed numbers. Sixth graders were selected for the study to ensure their lack of familiarity with the material because the topic of addition and subtraction of signed numbers is not taught until seventh or eighth grade in their school district.

Overall, students at the school ranked in the 90th percentile on statewide assessments of mathematics achievement, indicating that in spite of their lack of knowledge of signed arithmetic, students generally possessed good arithmetic skills. Fourteen lower achieving and 16 higher achieving students served in the SR group; 13 lower achieving and 17 higher achieving students served in the MR group. All students took a pretest and posttest consisting of easy and difficult problems, so comparisons of problem type are within-subject comparisons.

Materials

For each participant, the paper-and-pencil materials consisted of a pretest and a posttest. These tests were identical and consisted of a 8.5- × 11-in. sheet of paper containing 18 problems involving addition and subtraction of single-digit signed integers (listed in Table 1). Two problems of each of the following types were included: addition of two positive (P) numbers ($P + P$), addition of two negative (N) numbers ($N + N$), addition of a positive to a negative number ($N + P$), addition of a negative to a positive number ($P + N$), subtraction of two positive numbers ($P - P$), subtraction of two negative numbers ($N - N$), subtraction of a positive from a negative number ($N - P$), and subtraction of a negative from a positive number ($P - N$). Finally, two transfer problems in which one of the operands was the number zero were also included. The first eight problem types were classified as easy or difficult based on a median split of the proportion of correct answers given at pretest for each problem type (as listed in Table 1). The four easy problem types were $P + P$, $N + P$, $P + N$, and $P - P$; the four difficult problem types were $N + N$, $N - P$, $P - N$, and $N - N$.

For each participant, the computer-based materials consisted of a series of four 3.5-in., high-density floppy disks. Each disk contained an interactive computer program that corresponded to one of the four training sessions labeled with the training session date, the student's name, and the treatment group that the student belonged to (SR or MR). All disks contained two sets of 8 single-digit signed-arithmetic problems representing the same types as used on the tests, that is, ($P + P$), ($N + N$), ($N + P$), ($P + N$), ($P - P$), ($N - N$), ($N - P$), and ($P - N$). The first 8-item problem set for each training day was labeled as "Level 1" and the second set as "Level 2." Although the two sets of 8 problems were identified by different level numbers, the degree of difficulty of the problems was the same. Within each level, problem types were randomly ordered. The 16 problems contained in each of the four training sessions were identical for the SR and MR groups and are shown in Table 1. The computer program contained in the floppy disks was designed to create a personal log in the same disk that recorded all students' interactions during the training session (e.g., the typed in answers to the problems, the buttons clicked on the screen, and the exact time of each event). The programs were developed using Director 4.04 (Macromedia, 1994).

TABLE 1
Problems Used on Pretest–Posttest and Each Learning Session—Experiments 1 and 2

	Problem Type								
	A + *B* (Easy)	–*A* + *B* (Easy)	*A* + –*B* (Easy)	*A* – *B* (Easy)	–*A* + –*B* (Difficult)	–*A* – *B* (Difficult)	*A* – –*B* (Difficult)	–*A* – –*B* (Difficult)	Transfer
Proportion correct on pretest	.99	.65	.63	.62	.54	.34	.13	.41	.53
Pretest–posttest items	3 + 1 = 7 + 2 =	–4 + 2 = –6 + 9 =	2 + –9 = 6 + –3 =	5 – 7 = 2 – 5 =	–5 + –4 = –8 + –1 =	–2 – 3 = –3 – 1 =	1 – –8 = 7 – –2 =	–5 – –6 = –8 – –6 =	0 – –5 = –7 + 0 =
Session 1 items	2 + 5 = 2 + 3 =	–3 + 5 = –3 + 7 =	4 + –5 = 5 + –6 =	5 – 7 = 4 – 5 =	–4 + –3 = –2 + –3 =	–5 – 1 = –4 – 2 =	4 – –5 = 1 – –8 =	–2 – –6 = –1 – –7 =	
Session 2 items	4 + 1 = 6 + 2 =	–1 + 4 = –4 + 2 =	3 + –4 = 7 + –5 =	3 – 2 = 6 – 3 =	–5 + –4 = –6 + –2 =	–1 – 3 = –2 – 3 =	2 – –6 = 4 – –3 =	–9 – –6 = –7 – –5 =	
Session 3 items	1 + 7 = 8 + 1 =	–5 + 1 = –9 + 1 =	8 + –7 = 1 + –4 =	6 – 4 = 7 – 2 =	–1 + –8 = –3 + –3 =	–7 – 2 = –3 – 2 =	3 – –5 = 2 – –3 =	–6 – –2 = –5 – –6 =	
Session 4 items	4 + 2 =	–7 + 2 =	2 + –9 =	1 – 4 =	–2 + –5 =	–2 – 2 =	1 – –3 =	–3 – –6 =	

The apparatus consisted of 30 Macintosh computer systems that included either a 10-, 12- or 15-in. monitor and an average 40-megabyte internal hard disk. The computers were located in an elementary school's computer laboratory.

Procedure

One sixth-grade class was randomly assigned to be the treatment group, and the other sixth-grade class served as the control group. The distributions of student abilities and gender in the classes were equivalent as it is the school's policy to create equivalent classes at each grade level. Both classes had identical mathematics curricula and signed arithmetic was not covered in either class. Each student learned individually at a computer station during the training sessions so no whole-class instruction was involved in the study.

First, students were given a paper-and-pencil pretest during regular class time. Students who scored at or below the median (i.e., 50% correct) on the pretest were classified as low achieving and students who scored above the median were classified as high achieving.

Second, all students participated in each of the four training sessions held on different days over a 2 week period during regular class time. Each session was held in the school's computer laboratory, with each student seated at a Macintosh computer system. Students in the SR group were trained through a SR floppy disk, and students in the MR group were trained through a MR floppy disk. For each group, the disks were on computers in the school laboratory, prior to the student's arrival.

During each session for the SR group, students solved the 16 signed-arithmetic problems, working at their own rates in an interactive environment. First, as shown in the first frame of Figure 1, a main menu listing the first 8 problems in symbolic form (e.g., $4 - -5 =$ __) appeared on the screen. The learner could then select a problem by clicking on it. Then, as shown in the second frame of Figure 1, the selected problem appeared on the screen in symbolic form, prompting the learner to type in an answer. If the answer was correct, as shown in the third frame of Figure 1, the words "Yes! The answer is ___" appeared on the screen and learner could click on "Back to Menu" to go back to the main menu of problems. The fourth frame of Figure 1 shows the feedback given for a wrong answer: The words "Sorry! This is not the correct answer" appeared on the screen and the learner could either click on the button "Try Again" (to enter a new answer) or on the button "See Solution" (to be shown the correct answer). After presenting the correct answer, the program allowed the learner to move back to the main menu by pressing the "Back to Menu" button. Any time the student returned to the main menu, the list of the eight problems with a check mark added next to the completed problems appeared as shown in the fifth frame of Figure 1.

When given the main menu, the learner could select any of the eight problems, including those that were already checked. After selecting each of the first eight problems at least once, the learner could move on to the next set of problems by clicking on the "Done" button. The same procedure applied to the second set. After the learner completed the two sets of problems, the session ended.

Figure 2 shows selected frames from an example problem presented to the MR group. When a learner clicked on a problem, in addition to the symbolic form, a number line appeared with a bunny rabbit facing forward and standing on the 0 point (as shown in the first frame). Also, four buttons on the lower right part of the screen allowed the participant to move the bunny rabbit in four different combinations: face right and jump forward, face right and jump backwards, face left and jump forward, or face left and jump backwards. When the learner entered the correct answer, the words "Yes, the answer is __" appeared on the screen as in the SR group, and this was followed by an animated sequence that consisted of four major steps. For example, for the problem 4 – –5 = __, the 4 became highlighted and the words "FIRST: FIND MY STARTING POINT 4 means GO TO 4" appeared in a bubble above the bunny (in Frame 2); the bunny hopped to 4 on the number line (in Frame 3); the minus sign became highlighted and the words "SECOND: FIND THE OPERATION – means FACE LEFT" appeared in a bubble above the bunny (in Frame 4); the bunny turned to face left (in Frame 5); the –5 became highlighted and the words "THIRD: FIND HOW TO JUMP –5 MEANS JUMP BACK 5 STEPS" appeared in a bubble above the bunny (in Frame 6); the bunny hopped backwards five steps to 9 on the number line (in Frame 7); and the words "The answer is 9" appeared in a bubble above the bunny and the answer "9" appeared on the screen (in Frame 8). The rest of the procedure was identical to that of the SR group. After viewing the animation, the learner could either click on the button "See Solution Again" or on the button "Back to Menu." If the answer was wrong, the words "Sorry! This is not the correct answer" appeared on the screen and the learner could either click on the button "Try again" (to enter a new answer) or on the button "See Solution" (to be shown the correct answer and its respective animation). Any time the student returned to the main menu, the list of eight problems with a check mark added next to the completed problems would appear. After selecting each of the first eight problems at least once, the learner could move on to the next set of problems by clicking on the "Done" button. The same procedure applied to the second set. After the learner completed the two sets of problems, the session ended.

In all, both groups solved the same 64 problems, working independently, at their own rates and with the sole feedback of the computer program. The MR group received symbolic, visual, and verbal feedback, whereas the SR group received only symbolic feedback. After completing the four training sessions, all participants were given the paper-and-pencil posttest in their regular classrooms on a subsequent day.

Scoring

Pretest-to-posttest gain. For each student, we subtracted the number of correct answers on the pretest from the number of correct answers on the posttest to yield a pretest-to-posttest gain score.

Session-by-session learning score. For each student, we tallied the number of correct answers on each of the four training sessions to examine the gain from session to session.

Pretest-to-posttest change in arithmetic bugs. To better understand the cognitive processes involved during the students' training sessions, we also conducted an analysis of their strategies (Siegler & Jenkins, 1989) on problems in which they gave the wrong answer. The novelty of signed numbers for sixth-grade students is the ambiguity of the same symbol to refer to a subtraction operation and a negative number. Students that are in the process of learning the difference between the minus symbol as representing subtraction and as representing the sign of a number will have a tendency to confuse the two concepts. Consequently a "negative-bias effect" is expected to occur, in which the student treats the negative sign as a subtraction operator. Our analysis of arithmetic bugs applies only to problems containing an addition operator with two negatively signed numbers ($N + N$) or a subtraction operator and one negatively signed number ($N - P$, $P - N$). In our study, we assessed this bias by adding up for each student all the wrong answers where the negative sign in an operand was misinterpreted as an operator. For example, for the problem $-8 + -1 =$ ___, if a student gave as a result either a 7 or a -7, the answer was labeled as a *negative-bias bug*; an answer of 9 was labeled as a *conceptually good bug* because it did not reflect a negative bias. For the problem $-3 - 1 =$ ___, an answer of 2 or -2 indicated a negative-bias bug, whereas 4 indicated a conceptually good bug. For the problem $7 - -2 =$ ___, 5 or -5 reflected a negative-bias bug whereas -9 was a conceptually good bug. Thus, the good bug is not based on fundamental misconception of the meaning of a negative sign as a subtraction operator. Errors based on negative-bias interpretations suggest less understanding of the addition and subtraction of signed numbers than ones in which the operation is applied correctly.

RESULTS

Pretest-to-Posttest Gains

According to MR theory, the MR group should show a larger pretest-to-posttest gain than the SR group; cognitive-load theory makes the opposite prediction, and

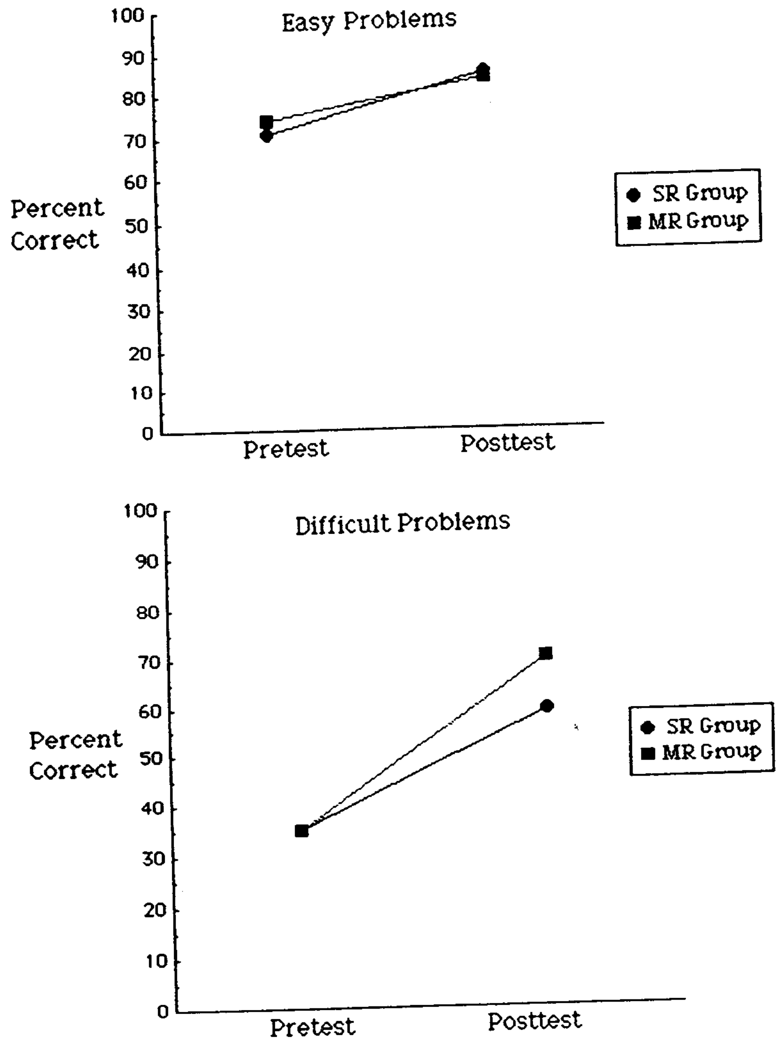

FIGURE 3 Proportion correct on pretest and posttest by single representation (SR) and multiple representation (MR) groups for difficult and easy problems—Experiment 1.

the hybrid theory predicts that the benefits of MR learning will be greatest for high-achieving students. Although the mean pretest-to-posttest gain by the MR group was greater than that of the SR group (*M*s = 4.10 and 2.97, *SD*s = 2.80 and 3.68, respectively), the difference failed to reach statistical significance, $F(1, 56) = 1.29$, $MSE = 10.32$, $p > .20$.

Given our suspicion that students had the most room for improvement on difficult problems (and may perform near the ceiling on easy problems), we reanalyzed the pretest-to-posttest gain data in a 2 × 2 analysis of variance (ANOVA) including treatment group (MR vs. SR) and problem difficulty (easy vs. hard) as factors. Figure 3 summarizes the scores in each treatment group for easy and difficult problems. For difficult problems, the mean pretest-to-posttest gain was 2.40 for the MR group (SD = 1.52) and 1.37 for the SR group (SD = 2.19); for easy problems the mean pretest to posttest gain was 1.17 for the MR group (SD = 1.70) and 1.30 for the SR group (SD = 2.07). An ANOVA revealed a significant interaction between group and problem difficulty, $F(1, 56) = 4.24$, $MSE = 10.21$, $p < .05$; supplemental Tukey tests with $p < .05$ confirmed that the MR group gained significantly more than the SR group for difficult problems, but the groups did not differ on easy problems.

Although these results tend to support MR theory, adding cognitive-load theory allows us to test an important prediction concerning the issue of for whom an MR experience would be productive. In particular, a hybrid theory predicts that the MR group will show a superiority over the SR group for high-achieving students but not for low-achieving students. Figure 4 summarizes the pretest-to-posttest gain scores for high- and low-achieving students in each treatment group. For high-achieving students, the mean gain for the MR group was 3.94 (SD = 3.07) and the mean gain for the SR group was .88 (SD = 2.42); for low-achieving students, the mean gain for the MR group was 4.31 (SD = 2.50) and the mean gain for the SR group was 5.69 (SD = 3.27). An ANOVA with group and achievement level as factors and gain as the dependent measure revealed a significant interaction between group and achievement level, $F(1, 56) = 12.33$, $MSE = 98.7$, $p < .01$; supplemental Tukey tests with $p < .05$ confirmed that the MR group gained significantly more than the SR group for higher achieving students, but the groups did not differ for lower achieving students.

We interpret the interaction as showing that the MR treatment produced better learning than the SR treatment for high-ability learners but not for low-ability learners. Low-ability learners, by definition, started at a low level of performance on signed-arithmetic problems; low-ability students showed equivalent improvements in the SR and MR groups, but mainly on easy problems and only up to the starting level of the high-ability learners. Given that the focus of the multimedia lesson was on understanding the complex problems, it is not surprising that the MR treatment was not particularly more effective than the SR treatment in promoting learning of the easy problems for the low-ability learners. In contrast, high-ability learners, by definition, started at a relatively higher level of performance—roughly equivalent to the final level of the low-ability learners; with room for improvement mainly in learning to solve difficult problems, the SR treatment resulted in no further improvement whereas the MR treatment produced substantial improvement. Consistent with the focus of the multimedia instruction in the MR treatment, the MR treatment was particularly effective for students who had already mastered the basics of signed addition with easy problems. Thus, although three groups show pretest-to-posttest improvements, the improvements of low-ability students in the SR and MR groups reflect starting nearer to the "floor" and mainly learning to solve easy problems, whereas the improvements of the high-ability students in the MR group reflect starting nearer to the "ceiling" and mainly learning to solve difficult problems.

Session-by-Session Learning

According to MR theory, the MR group should learn faster than the SR group, whereas cognitive-load theory makes the opposite prediction. Finally, by combining the two theories we can predict that the MR group will learn faster than the SR group for high-achieving students but not for low-achieving students. The left panel of Figure 5 summarizes the scores on each session in each treatment group. For Ses-

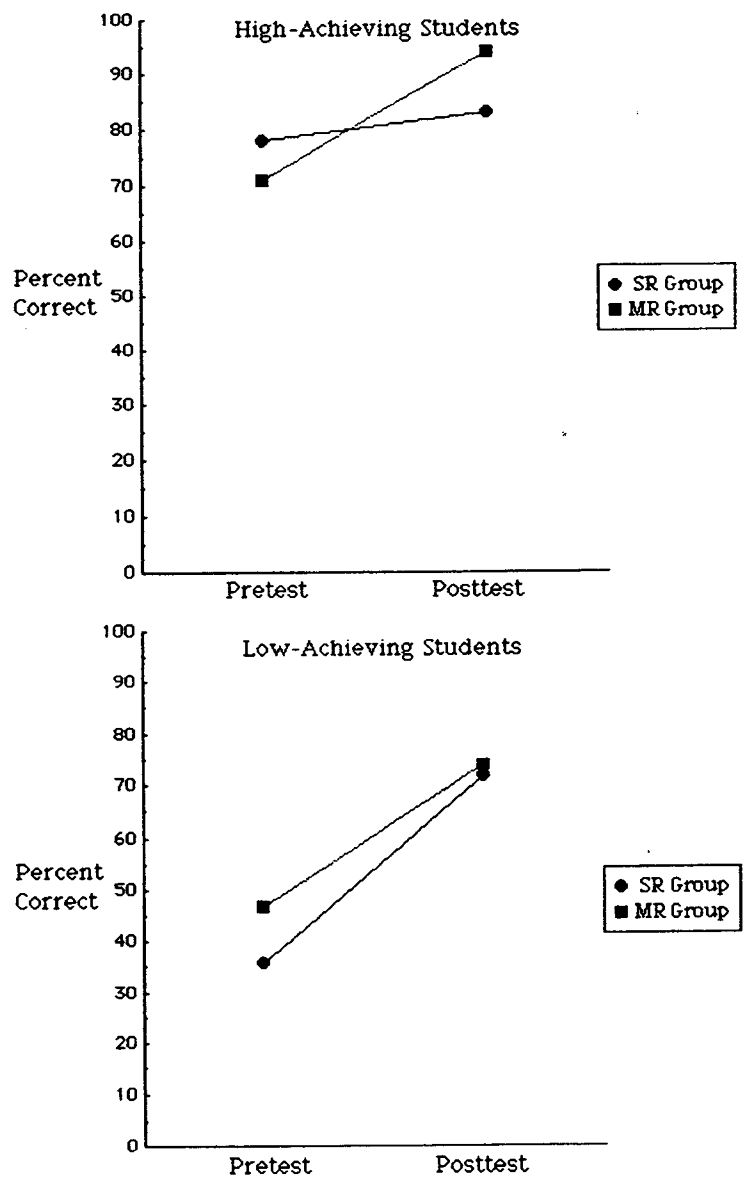

FIGURE 4 Proportion correct on pretest and posttest by single representation (SR) and multiple representation (MR) groups for high-achieving and low-achieving students—Experiment 1.

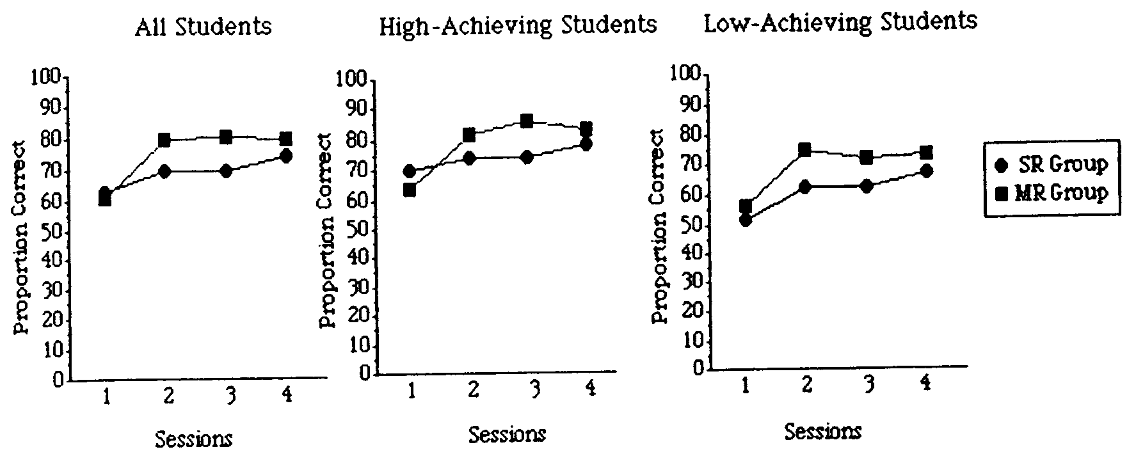

FIGURE 5 Proportion correct on each learning session by single representation (SR) and multiple representation (MR) groups for all, high-achieving, and low-achieving students–Experiment 1.

sions 1 to 4, the mean number of correct answers on each session was 9.96, 11.07, 11.10, and 11.77, respectively for the SR group (*SD*s = 3.31, 2.63, 2.67, and 3.13, respectively); and 9.70, 12.63, 12.83, and 12.70 for the MR group (*SD*s = 2.49, 2.31, 2.18, and 2.82, respectively). An ANOVA with group (SR vs. MR) as a between-subjects factor and training session as a within-subject factor revealed a significant interaction, $F(3, 168) = 4.09$, $MSE = 10.78$, $p < .01$, consistent with the observation that students in the MR group learned significantly faster than students in the SR group. These results are consistent with the predictions of the MR hypothesis.

Although MR theory is consistent with the foregoing results, cognitive-load theory suggests an important additional prediction concerning individual differences in the effectiveness of MR training. In particular, a hybrid theory (i.e., based on MR theory and cognitive-load theory) predicts that the MR group will learn faster than the SR group for high-achieving students but not for low-achieving students. Therefore, we conducted an ANOVA for each ability group separately on the session-by-session scores with group as between-subjects factor and session as a within-subjects factor. The middle panel of Figure 5 summarizes the scores on each learning session for high-achieving students in each treatment group. For high-achieving students, the mean scores on sessions 1 through 4 were 11.24, 11.88, 11.88, and 12.53, respectively, for the SR group (*SD*s = 3.29, 2.29, 2.34, and 2.85, respectively) and 10.18, 13.12, 13.77, and 13.35, respectively, for the MR group (SDs = 2.24, 1.62, 1.64, and 2.85, respectively). The right panel of Figure 5 summarizes the scores on each learning session for low-achieving students in each treatment group. For low-achieving students, the mean scores on Sessions 1 through 4 were 8.31, 10.00, 10.08, and 10.77, respectively, for the SR group (SDs = 2.59, 2.74, 2.81, and 3.30, respectively) and 9.08, 12.00, 11.62, and 11.85, respectively, for the MR group (SDs = 2.75, 2.94, 2.26, and 2.64, respectively). The ANOVA conducted on the session-by-session scores for high-achieving students revealed a significant interaction between group and session, $F(3, 96) = 6.09$, $MSE = 13.58$, $p < .01$, consistent with the interpretation that MR students learned faster than SR students. In contrast, the ANOVA conducted on the session-by-session scores for low-achieving students failed to produce a significant Group × Session interaction, $F(3, 72) = .59$, $MSE = 1.89$, $p > .50$, suggesting that the groups did not differ in their pattern of improvement. Overall, these results are most consistent with a combination of MR and cognitive-load theories.

Pretest-to-Posttest Changes in Arithmetic Bugs

MR theory predicts that a comparison of the type of errors at pretest and posttest would show that students in the MR group learned to choose better strategies than students in the SR group. Figure 6 summarizes the number of negative-bias bugs on pretest and posttest in each treatment group. At pretest, the mean number of negatively biased answers was 1.40 ($SD = 1.45$) for the MR group and 0.90 ($SD = 0.85$) for the SR group; at posttest, the MR mean number of negatively biased answers was reduced to 0.53 ($SD = 0.68$), whereas the SR mean number of negatively biased answers was unaltered at 0.90 ($SD = 1.24$). An ANOVA was conducted with group (SR vs. MR) as the between-subjects factor, test (pretest vs. posttest) as a within-subjects factor, and number of negatively biased answers as the dependent measure. The ANOVA provided evidence for a significant interaction between group and test, $F(1, 58) = 4.97$, $MSE = 5.63$, $p < .05$. Consistent with the MR theory, students in the MR group reduced their negative bias from pretest to posttest significantly more than did students in the SR group.

EXPERIMENT 2

Consistent with MR theory, the results of Experiment 1 showed an advantage of MR instruction over SR instruction across three different dependent measures: The MR group learned more (on difficult problems), learned faster, and reduced more of their negative-bias bugs than did the SR group. An important additional result consistent with cognitive-load theory is that, in general, these effects were strongest for high-achieving students. Overall, the results of Experiment 1 are consistent with a hybrid theory in which learning with MRs is more effective than learning

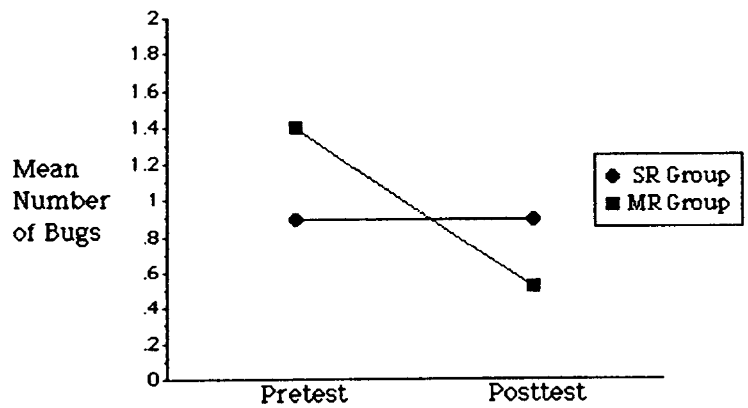

FIGURE 6 Mean number of negative-bias bugs on pretest and posttest by single representation (SR) and multiple representation (MR) groups—Experiment 1.

with an SR mainly when the learner's working memory is not overloaded by the instructional presentation. We conducted a second experiment to examine in more detail the role of cognitive load in learning with MRs. In particular, we examined the effectiveness of MR training (as in Experiment 1) for students who scored high or low in spatial ability and for students who scored high or low in memory span.

According to cognitive-load theory, low spatial ability students may be less able than high spatial ability students to take advantage of the presentation of MRs. When visual and verbal material are presented contiguously, cognitive load results from the need to hold and manipulate the visual representation in working memory while making the respective integration with the verbal representation (Mayer & Sims, 1994). Low spatial students must devote many cognitive resources to holding the images in working memory, leaving inadequate cognitive resources for the process of integrating the images with verbal representations. In contrast, high spatial ability students can hold images in working memory without expending many cognitive resources, allowing them to use their cognitive resources for the process of building connections among representations. Although spatial ability has been defined in several ways (Carroll, 1993; Lohman, Pellegrino, Alderton, & Regian, 1987), we have focused on spatial visualization in our study. This aspect of spatial ability is most relevant to learning from complex animations and can be measured by testing students' abilities to mentally rotate or fold two- or three-dimensional objects (Mayer, 1997). Based on the hybrid theory combining cognitive-load and MR theory, we predicted that high spatial students would benefit more from our MR treatment than would low spatial students.

According to cognitive-load theory, working memory capacity is limited in both duration and capacity (Miller, 1956; Simon, 1974). When students are required to solve problems that exceed their working memory capacity, learning may be hampered (Sweller, 1988, 1989). According to capacity theories of comprehension (Just & Carpenter, 1992), capacity constrains comprehension, and it does more for some individuals than for others. Students with larger capacity are better able to maintain and coordinate MRs in working memory than are students with low capacities. Congruent with this view, we predicted that students with high working memory span would be better able to take advantage of learning in high cognitive-load conditions imposed by our MR presentation than would low working memory students. The overarching goal of Experiment 2 is to examine how individual differences in spatial ability and working memory capacity affect learning with MRs in a multimedia environment.

METHOD

Participants and Design[1]

The participants were 26 sixth-grade students from the same population as Experiment 1. All students took a pretest and posttest and participated in four training sessions, so pretest-to-posttest comparisons and session-by-session comparisons are within-subject comparisons.

Materials

For each participant, the paper-and-pencil materials consisted of a pretest, a battery of four cognitive tests, and a posttest. The pretest and posttest were identical to those described for Experiment 1. The cognitive tests were Part 1 of the Paper Folding test, Part 1 of the Cube Comparisons test, the Auditory Letter Span test, and the Number Comparison test from the Kit of Factor-Referenced Cognitive Tests distributed by the Educational Testing Service (Ekstrom, French, & Harman, 1976). In the Paper Folding test, the student must imagine that a sheet of paper has been folded in a certain way, a hole is punched through all thicknesses of the paper at a certain point, and then the sheet is unfolded. The folding and punching are indicated on the left side of the sheet, and the student must select which of five unfolded sheets is the result. In the Cube Comparisons test, the student must tell whether two cubes are the same or different; if the cubes match if one were rotated, then the correct answer is "same"; however, if they don't, then the correct answer is "different." In the Auditory Letter Span test, students are read 15 numbers, one by one, digit by

[1]Experiment 2 originally included a second group of sixth graders who received the SR treatment as in Experiment 1. However, the teacher provided an in-class unit on addition and subtraction of signed numbers immediately prior to the study, so data for this group could not be used. All other students in Experiment 1 and 2 had no prior training in addition and subtraction of signed numbers.

digit. Immediately after each number is read, the student has to write down as many digits as he or she can remember. The number of digits in the numbers vary from 3 to 13. Finally, in the Number Comparison test, the student is presented with pairs of numbers and must decide whether or not the two numbers are the same or not by putting an *X* between them when they are different. For example, for the pair 73845 __ 73855, the student should put an *X* on the line between the numbers. Each test was printed on an 8.5- × 11-in. sheet of paper. For each participant, the computer-based materials consisted of a series of four 3.5-in., high-density floppy disks identical to those used for the MR group in Experiment 1. The apparatus was the same as in Experiment 1.

Design and Procedure

First, students were given a paper-and-pencil pretest during regular class time. Students who scored at or below the median (i.e., 50% correct) on the pretest were classified as low achieving, and students who scored above the median were classified as high achieving.

Second, students took the battery of four cognitive tests, which were administered following the pretest. For each test, students read instructions that included a worked-out sample problem. The first test was the Auditory Letter Span test. After students read the instructions, they were read the numbers from the test. Immediately after each number was read, the students had to write down as many digits as they could remember. Once done with the last number, students were instructed to put their pencils on the desk and wait until the experimenter collected all the test sheets. The next tests were the Number Comparison test, the Cube Comparisons test, and the Paper Folding test, administered, respectively, in this order. For each test, after students had read the instructions, they were given 3 min to complete as many test items as possible. When the 3 min were over, students were instructed to stop writing, put their pencils on the desk, and wait until the experimenter collected all the test sheets.

Third, all students participated in each of the four training sessions held on different days over a 2-week period during regular class time, as in Experiment 1.

Finally, after completing the four training sessions, all students were given the paper-and-pencil posttest in their regular classrooms on a subsequent day.

Scoring

Pretest-to-posttest gain. For each student, we subtracted the number of correct answers on the pretest from the number of correct answers on the posttest to yield a pretest-to-posttest gain score.

Spatial ability score. For each student, we tallied the number of correct answers on the Cube Comparisons test and on the Paper Folding test. After standardizing each score, a composite spatial ability score was computed for each participant by adding both scores. For purposes of analysis, 11 students who scored above the mean were classified as high spatial ability, and 15 who scored at or below the median were classified as low spatial ability.

Memory span score. For each student, we tallied the number of correct answers on the Auditory Letter Span test and on the Number Comparison test. After standardizing each score, a composite working memory span score was computed for each participant by adding both scores. For purposes of analysis, 11 students who scored above the mean were classified as high memory span, and 15 students who scored at or below the mean were classified as low memory span.

RESULTS

- *Do high spatial ability students who receive MR training learn addition and subtraction of signed numbers better than low spatial ability students who receive MR training?* Based on a hybrid theory combining cognitive-load and MR theories, we predicted that high spatial ability students who trained with MR would outperform low spatial ability students who trained with MR. Consistent with this prediction, the high spatial ability students produced a significantly greater pretest-to-posttest gain ($M = 4.46$, $SD = 3.24$) than did the low spatial ability students ($M = 0.67$, $SD = 4.73$), $t(24) = 2.29$, $p < .05$. Our prediction that high spatial ability students in the MR group would outperform low spatial ability students was confirmed.
- *Do high working-memory span students who receive MR training learn addition and subtraction of signed numbers better than low working-memory span students who receive MR training?* Also congruent with a hybrid theory combining cognitive-load and MR theory, we predicted that high working-memory span students who learned with MRs in a multimedia environment would outperform low working-memory span students who learned in the same way. Although the mean pretest-to-posttest gain by the high working-memory span group was greater than the mean pretest-to-posttest gain of the low working-memory span group (Ms = 3.55 and 1.33, SDs = 4.03 and 4.74, respectively), the difference failed to reach statistical significance, $t(24) = 1.25$, $p > .20$. One explanation for the failure to confirm our prediction is that our measure of working memory capacity may have been inadequate. Most of the past research has focused on working memory as the storage of information for retrieval after a brief interval. The two tests that we used to measure memory span are congruent with this view. However, more recently the view

of working memory has been broadened to include not just the storage of items for later retrieval but also the storage of partial results in complex sequence computations such as language comprehension (Just & Carpenter, 1992). Our study required students to quickly process text, symbols, and an animation. In this framework, a better way to assess working memory capacity might be to use listening or reading span tasks, devised to simultaneously draw on the processing and storage resources of working memory (Daneman & Carpenter, 1980).

Overall, the results of Experiment 2 provide important new evidence that experience in an MR treatment is more effective for students with high spatial ability than for students with low spatial ability. In short, Experiment 2 clarifies for whom MRs are most helpful and helps to show how instruction with MRs is influenced by the ability of learners to handle the cognitive load imposed by processing information in multiple forms.

DISCUSSION

The goal of this study was to explore the cognitive implications of training students with MRs in a high-cognitive load situation. MR theory proposes that when solving an arithmetic problem, instructional methods that promote the use of MRs (i.e., symbolic, visual, and verbal) are more likely to aid in the learning process than instructional methods that rely on SRs (i.e., symbolic). Cognitive-load theory claims that when learners are required to mentally integrate disparate sources of mutually referring information such as text, symbols, and animations, such split-source information may create a heavy cognitive load that disrupts learning. Therefore, when using multimedia instructional methods we are faced with a trade-off problem. On one hand, multimedia can facilitate learning by representing concepts in more than one modality. On the other hand and congruent with cognitive-load theory, the more representations offered for the same procedure, the heavier the cognitive load and the harder the learning. Our study provided evidence that learning in a multimedia MR environment benefits higher achieving and high spatial ability students the most, although lower achieving and low spatial ability students do not learn significantly more from MRs than from an SR. These findings have important theoretical and practical implications.

The significantly larger gain with MR over SR for the high-achieving group of students can be theoretically accounted for by integrating cognitive-load and MR theories. Given that the amount of resources in working memory is limited, and given that the high-achieving group had already made some information automated (they were mastering relatively better the symbolic representation for the problems at pretest), they should consume less of the memory resources than the low-achieving group. Thus, they are able to benefit from MR presentations that require more cognitive load than SR presentations. Comparatively, the low-achieving group had to use more resources in working memory to build the symbolic representation than the high-achieving group. Because their cognitive load was high even with the SR treatment, lower achieving students are likely to experience overload in the MR treatment. As a result, adding extra visual and verbal explanations did not help them. Moreover, the results suggest that lower achieving and low spatial ability students are either ignoring, or unable to process, the additional information. This analysis shows how two seemingly competing theories (MR and cognitive load) can be reconciled.

It should also be noted that cognitive-load theory does not predict that the additional information is harmful per se. In a recent review, Sweller and his colleagues (Sweller, van Merrienboer, & Paas, 1998) addressed some instructional design issues related to worked examples. For example, it is important to determine if the additional information that is provided in a worked example is redundant or if it is essential for understanding the primary source of information. In worked examples with multiple sources of redundant information, the additional information can have negative consequences by forcing the students to split their attention between them. On the other hand, when worked examples present extra sources of information that are needed for the lesson to be intelligible, then the additional information can have positive consequences (Sweller, 1993). In our studies, the MR groups were given worked examples with full verbal and visual explanations that made explicit the procedure of how to solve an arithmetic problem successfully. The SR groups, instead, were not provided with any type of explanation. Although cognitive-load theory would predict better learning from nonredundant worked examples than from conventional problem-solving practice, it was found in this study that only the higher achieving and high spatial ability students did benefit from the explanations. These results suggest that even when the additional explanations can be helpful to advance students' learning by providing detailed worked examples, students need to be equipped with the necessary cognitive resources to process the explanation.

The strategy analysis showed a strong negative sign bias for both groups at pretest. Despite this fact, evidence was found that providing an MR model to solve problems allowed the treatment group to significantly lower the bias at posttest, whereas providing an SR model left the bias intact for the control group. A possible interpretation of these results can be made through mental models theory (Johnson-Laird, 1983). This theory claims that in the process of creating a mental model, people might fail due to limitations in working memory capacity or to the influence of prior knowledge. In the context of our study, sixth-grade students have already built a model to solve addition and subtraction of natural numbers. Therefore, the original negative bias at pretest might be caused by trying to interpret the new model for signed numbers through their prior knowledge of natural numbers. That is, a negative sign preceding a number is interpreted according to its meaning in the old model that is subtraction. Providing an MR model to the treatment group proved to help make their bugs better. After the training sessions, even for the problems that were not yet learned, the MR group used significantly better strategies than the SR group. Their errors were at least showing an understanding of the difference between the subtraction operation and the negative sign of a number. This suggests that MR training not only produced quantitative differences in performance (measured by the larger mean gain score) but also produced qualitative differences in their learning. Students in the MR group are in the process of building a successful model to add and subtract integers. Students in the SR group are instead making the same type of errors that they had before training.

Experiment 2 provided evidence for an advantage to learn from MRs for higher spatial ability students. Given that working memory appears to be a limiting factor in the integration of multiple sources of mutually referring information, this result is not surprising. Individuals who are better able to hold and manipulate visual representations in memory will especially benefit when the mapping of text, symbols, and complex animations is required. Our results then, confirm the capacity theories of comprehension in the area of mathematics learning with multimedia.

On the practical side, this study provides results with direct instructional implications. First, it provides empirical support for using MR's to help students learn mathematical procedures. When designing educational software, an interactive computer-based learning environment enables better learning when it includes symbolic, visual, and verbal representations that are coordinated in time rather than solely symbolic representations.

Second, the benefits of using MRs for example problems are strongest on difficult problems and for students who already have a good knowledge of the basic arithmetic of natural numbers. This suggests that in the teaching of mathematics, the ideal learning situation might be to first bring the lower achieving students to a higher level of proficiency on the symbolic representations so that they can benefit more fully from the multimedia, multirepresentational environment.

Third, for instructional material to be effective, it must consider the role that individual differences can play in the learning process, especially if learners are required to integrate multiple sources of mutually referring information. To understand materials that are high in element interactivity, such as mathematics, it is important to design the materials in a manner that minimizes cognitive load. A high cognitive-load environment may have positive learning effects only when students possess the necessary cognitive resources to process simultaneously in working memory the MRs. Spatial ability is a potential indicator of successful learning in such environments. Other indicators, such as reading span, should be analyzed in future research.

Finally, more research is needed on learning from multimedia environments. It is possible that some of these results do not generalize to other domains or to materials that are low in interactivity. Future research should extend this study to new areas requiring mental integration between disparate sources of information.

ACKNOWLEDGMENTS

We thank Kathy Gerber, Beth Thompson, and Bianca Jamgochian, teachers at the Mountain View Elementary School, Goleta, California, for kindly participating in this experiment.

REFERENCES

Alexander, P. A., White, C. S., & Daugherty, M. (1997). Analogical reasoning and early mathematics learning. In L. English (Ed.), *Mathematical reasoning: Analogies metaphors, and images* (pp. 117–147). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Baddeley, A. (1992). Working memory. *Science, 255,* 556–559.

Bassok, M. (1997). Two types of reliance on correlations between content and structure in reasoning about word problems. In L. English (Ed.), *Mathematical reasoning: Analogies metaphors, and images* (pp. 221–246). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Brenner, M., Mayer, R. E., Mosely, B., Brar, T., Duran, R., Reed, B. S., & Webb, D. (1997). Learning by understanding: The role of multiple representations in learning algebra. *American Educational Research Journal, 34,* 663–690.

Brownell, W. A. (1928). *The development of children's number ideas in the primary grades.* Chicago: University of Chicago Press.

Brownell, W. A. (1935). Psychological considerations in the learning and teaching of arithmetic. In W. D. Reeve (Ed.), *The teaching of arithmetic: The tenth yearbook of the National Council of Teachers of Mathematics* (pp. 1–31). New York: Teachers College Press.

Bruer, J. T. (1993). *Schools for thought.* Cambridge, MA: MIT Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, England: Cambridge University Press.

Case, R., & Okamoto, Y. (Eds.). (1996). The role of central conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development, 61*(1–2).

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8,* 293–332.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3,* 149–210.

Cooper, G., & Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79,* 347–362.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Davis, R. B., & Maher, C. A. (1997). How students think: The role of representations. In L. D. English (Ed.), *Mathematical reasoning: Analogies metaphors, and images* (pp. 93–115). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests.* Princeton, NJ: Educational Testing Service.

English, L. D. (Ed.). (1997). *Mathematical reasoning: Analogies, metaphors, and images.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Fuson, K. (1992a). Research on learning and teaching addition and subtraction of whole numbers. In G. Leinhardt, R. Putnam, & R. Hattrup (Eds.), *Analysis of arithmetic for mathematics teaching* (pp. 53–187). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Fuson, K. (1992b). Research on whole number addition and subtraction. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 243–275). New York: Macmillan.

Gentner, D., & Stevens, A. L. (Eds.). (1983). *Mental models.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Glynn, S., Duit, R., & Thiele, R. (1995). Teaching science with analogies: A strategy for constructing knowledge. In S. Glynn & R. Duit (Eds.), *Learning science in schools: Research reforming practice* (pp. 247–243). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Goswami, U. (1992). *Analogical reasoning in children.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Griffin, S., & Case, R. (1996). Evaluating the breadth and depth of training effects when central conceptual structures are taught. *Monographs of the Society for Research in Child Development, 61*(1–2), 83–102.

Griffin, S., Case, R., & Capodilupo, A. (1995). Teaching for understanding: The importance of central conceptual structures in the elementary mathematics curriculum. In A. McKeough, J. Lupart, & A. Marini (Eds.), *Teaching for transfer: Fostering generalization in learning* (pp. 123–151). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Grouws, D. A. (Ed.). (1992). *Handbook of research on mathematics teaching and learning* (pp. 243–275). New York: Macmillan.

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

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65–97). New York: Macmillan.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, England: Cambridge University Press.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122–149.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving arithmetic word problems. *Psychological Review, 92,* 109–129.

Kozma, R. B. (1991). Learning with media. *Review of Educational Research, 61*, 179–211.

Kozma, R. B., Russell, J., Jones, T., & Marx, N. (1996). The use of multiple, linked representations to facilitate science understanding. In S. Vosniadou, E. DeCorte, R. Glaser, & H. Mandl (Eds.), *International perspectives on the design of technology-supported learning environments* (pp. 41–60). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lakoff, G., & Johnson, D. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Lakoff, G., & Nunez, R. E. (1997). The metaphorical structure of mathematics: Sketching out cognitive foundations for a mind-based mathematics. In L. D. English (Ed.), *Mathematical reasoning: Analogies metaphors, and images* (pp. 21–89). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lewis, A. B. (1989). Training students to represent arithmetic word problems. *Journal of Educational Psychology, 81,* 521–531.

Lohman, D. F., Pellegrino, J. W., Alderton, D. L., & Regian, J. W. (1987), Dimensions and components of individual differences in spatial abilities. In S. H. Irvine & S. E. Newstead (Eds.), *Intelligence and cognition: Contemporary frames of reference* (pp. 253–312). Dordrecht, The Netherlands: Martinis Nijhoff.

Macromedia [Computer software]. (1994). *Director 4.0.* San Francisco: Author.

Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology, 88,* 49–63.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32,* 1–19.
Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-dual coding hypothesis. *Journal of Educational Psychology, 83,* 484–490.
Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86,* 389–401.
Mayer, R. E., Sims, V. K., & Tajika, H. (1995). A comparison of how textbooks teach mathematical problem solving in Japan and the United States. *American Educational Research Journal, 32,* 443–460.
Miller, G. (1956). The magical number 7 plus or minus two. Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.
Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice Hall.
Nunes, T., Schliemann, A. D., & Carraher, D. W. (1993). *Street mathematics and school mathematics.* New York: Cambridge University Press.
Ortony, A. (Ed.). (1993). *Metaphor and thought (2nd ed).* New York: Cambridge University Press.
Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology, 84,* 429–434.
Paivio, A. (1986). *Mental representations: A dual coding approach.* Oxford, England: Oxford University Press.
Piaget, J. (1952). *The origins of intelligence in children.* New York: Norton.
Resnick, L. B. (1983). A developmental theory of number understanding. In H. P. Ginsberg (Ed.), *The development of mathematical thinking* (pp. 109–152). New York: Academic.
Resnick, L. B., & Ford, W. W. (1981). *The psychology mathematics for instruction.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Simon, H. A. (1974). How big is a chunk? *Science, 183,* 482–488.
Sowder, J. T. (1992). Making sense of numbers in school mathematics. In G. Leinhardt, R. Putnam, & R. Hattrup (Eds.), *Analysis of arithmetic for mathematics teaching* (pp. 1–51). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12,* 257–285.
Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology, 81,* 457–466.
Sweller, J. (1993). Some cognitive processes and their consequences for the organization and presentation of information. *Australian Journal of Psychology, 45,* 1–8.
Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in structuring technical material. *Journal of Experimental Psychology: General, 119,* 176–192.
Sweller, J., van Merrienboer, J. G., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10,* 251–296.
Thorndike, E. L. (1913). *Educational psychology.* New York: Teachers College Press.
Trafton, J. G., & Reiser, B. J. (1993). The contribution of studying examples and solving problems to skill acquisition. In *Proceedings of the 15th annual conference of the Cognitive Science Society* (pp. 1017–1022). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Vosniadou, S., & Ortony, A. (Eds.). (1989). *Similarity and analogical reasoning.* Cambridge, England: Cambridge University Press.
Wertheimer, M. (1959). *Productive thinking.* New York: Harper & Row.
Zhu, X., & Simon, H. (1987). Learning mathematics from examples by doing. *Cognition and Instruction, 4,* 137–166.