Do Multiple Representations Need Explanations? The Role of Verbal Guidance and Individual Differences in Multimedia Mathematics Learning

Roxana Moreno
University of New Mexico

Richard Durán
University of California, Santa Barbara

Elementary school children, some of whom were nonnative speakers of English, learned to add and subtract integers in a discovery-based multimedia game either with or without verbal guidance in English or optionally in Spanish (Groups G—verbal guidance and No-G—no verbal guidance, respectively). Group G members chose to listen to verbal explanations in their first language and showed larger posttest scores than Group No-G. High-computer-experience students in Group G outperformed the rest of the students on training session scores and a transfer test. Longer latencies to respond to practice problems affected all learning measures positively. Results support the use of verbal guidance for discovery-based multimedia games and show that multimedia games may not be equally effective for all learners.

What is the role of verbal guidance in promoting mathematics learning from a discovery-based multimedia game? Are there important individual differences for which a multimedia game helps some kinds of learners more than others? These are important questions both for research and for the application of research, to improve instruction and learning outcomes in learning mathematics with multimedia programs. Our first question is concerned with the issue of how one should design multimedia games to enhance student learning and focuses on the role of verbal guidance as an aid to meaning making. To answer this question, we examined the cognitive consequences of learning how to add and subtract positive and negative numbers within a discovery-based multimedia game either with or without verbal guidance. Our second question is concerned with the issue of how individual differences influence learning from interactive multimedia games. More specifically, we were interested in examining the role that students' computer experience and latency to solve problems had when students were asked to learn from a discovery-based computer game that presented multiple representations (symbolic and visual) of the arithmetic procedure.

The main goal of this study was to investigate if the benefits of guidance that had been documented in past research on discovery-based strategies in the classroom extended into the realm of multimedia learning. In our study, elementary students studied 16 addition and subtraction problems involving positive and negative numbers in each of four training sessions with two representations of the arithmetic procedure: a symbolic representation consisting

of the number sentence (i.e., $2 - {-5} = \underline{\quad}$) and a visual representation using a number line and a computer animation to show how the symbolic number sentence relates to a bunny's movements along the number line. One group of students learned without verbal guidance and was presented with the symbolic and visual representations alone (Group No-G). Another group of students learned with identical symbolic and visual representations but was presented, additionally, with a verbal description (in English or optionally in Spanish) explaining how the symbols related to the bunny's movements along the number line (Group G). We measured learning in three ways: pretest to posttest improvement 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 the transfer of the arithmetic procedure to solving word problems.

By addressing the role of verbal guidance and individual differences in learning with discovery-based multimedia, educational psychologists can develop a more robust scientific foundation for theories of multimedia learning and instruction and provide a stronger knowledge base for informing instructional practice of value to particular students. We begin this inquiry by describing in detail the discovery-based learning environments used in our study. Then, we explain the case for using multiple representations and discovery-based methods in multimedia learning. Finally, we present a cognitive theory of multimedia learning (CTML) to help derive predictions of the effectiveness of our treatments for each one of the individual differences of interest.

A Discovery-Based Multiple-Representation Environment

Figure 1 shows selected frames from an example problem presented to students in Group No-G. First, students selected one of eight problems to solve from the problem menu (as shown in the first frame of Figure 1). Then, as shown in the second frame of Figure 1, students saw the problem presented in symbolic form (as $4 - {-5} = \underline{\quad}$) and a number line showing integers from -9 to 9, with a bunny standing at the 0 point. A simulated joystick con-

4 - -5 =
-5 - 1 =
-2 - -6 =
-4 + -3 =
2 + 5 =
-3 + 5 =
4 + -5 =
5 - 7 =

LEVEL ONE

4 - -5 =

-9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9

4 - -5 =

-9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9

4 - -5 =

-9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9

4 - -5 =

-9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9

4 - -5 =

-9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9

Figure 1. Selected frames from the no-guidance program.

sisting of the following four alternative moves for the bunny appeared in the lower right corner: face to the left, face to the right, jump forward one step, or jump backward one step. Students could click on any combination of the four joystick options and instantly see the resulting change in the bunny on the number line. The program instructed learners to try to figure out the problem by moving the bunny along the number line using the joystick. When students were ready to answer, they typed in a numeral (and negative sign, if needed).

If the student's answer was not correct, the student could try again or see the solution to the problem. If students typed in the correct answer, they heard the word "yes," which was followed by an animated sequence that consisted of four major steps in solving the problem. First (in the third frame of Figure 1), the symbol 4 was highlighted, and the bunny moved to position 4 on the number line. Thus, students saw how the first step in the procedure could be represented in symbols and movements. Second (in the fourth frame of Figure 1), the minus sign was highlighted, and the bunny turned to face the left side of the screen. Third (in the fifth frame in Figure 1), the symbol −5 was highlighted, and the bunny made five jumps to the right. Finally (in the sixth frame of Figure 1), the number 9 was highlighted, and the bunny faced forward on the ninth position of the number line. At this point, the student could click on *See Solution Again,* which would repeat this four-step animated sequence, or on *Back to Menu,* which would take the student back to the menu frame shown in the first frame of Figure 1.

The multimedia game presented to students in Group G was identical to that of Group No-G with one exception. For every problem, the learner could also hear a guided explanation describing each step of the respective animation. For the above example problem $4 − −5 = \_\_$, the student heard the words "First, find my starting point: Four means go to 4" as the symbol 4 was highlighted, followed by the image of the bunny moving to position 4 on the number line. Thus, the learner saw how the first step in the procedure could be represented in symbols, words, and pictures. Second, while the minus sign was highlighted, the student could hear the words "Second, find the operation: To subtract means face left" and then watch the bunny turn to face the left side of the screen. Third, while the symbol −5 was highlighted, the student could hear the words "Third, find the direction to jump: Negative five means jump back five steps" and then watch the bunny make five jumps to the right. Finally, while the number 9 in the answer box was highlighted, the student could hear the words "The answer is 9" and watch how the bunny turned to face forward on the 9 position in the number line. The guided version also contained a button labeled *Español* on the lower right corner of the computer screen. By clicking on this button, students heard a Spanish translation of the English explanation.

The Case for Multiple Representations

The case for using multiple representations of knowledge in instruction is based on the idea that teaching with more representations facilitates and strengthens the learning process by providing several mutually referring sources of information (Grouws, 1992; Kozma, Russell, Jones, & Marx, 1996; National Council of Teachers of Mathematics, 1989; Wagner & Kieran, 1989). Representing concepts or procedures in more than one format, which is

called providing multiple knowledge representations (Brenner et al., 1997; Spiro, Feltovich, Jacobson, & Coulson, 1992), allows learners to construct understandings that prepare them better for transfer, with each example and representation adding connections and perspectives that others miss (Sternberg & Frensch, 1993).

The multimedia game used in the present study combined symbolic and visual representations of the arithmetic procedure of addition and subtraction of integers in an effort to help students build connections between the formal computational procedures that use symbols (the number sentence) and their informal conceptual knowledge about moving along a path (the movements of a bunny along the number line). Adding a visual representation of an arithmetic procedure to the more traditional symbolic representation can be particularly appropriate for less skilled students who lack formal academic training in the subject domain by building on their existing intuitive knowledge (English, 1997; Fuson, 1992a, 1992b; Halford, 1993; Hiebert & Carpenter, 1992; Kintsch & Greeno, 1985). We chose the number line for a visual representation because it has been implicated as a central conceptual structure underlying number sense (Case & Okamoto, 1996; Griffin & Case, 1996), as a concrete manipulative for understanding arithmetic (Brenner et al., 1997; Hiebert & Carpenter, 1992; Lewis, 1989; Moreno & Mayer, 1999b), and as a grounding metaphor for arithmetic (Lakoff & Nunez, 1997). According to the walking-along-a-path 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" (Lakoff & Nunez, 1997, p. 37).

The Case for Discovery-Based Multimedia Environments

Discovery-based learning environments are intended to facilitate learning through creative problem-solving experiences that de-emphasize formal didactic instruction (Bruner, 1961; Wittrock, 1966). Gagné (1965) distinguished between two kinds of learning on the basis of the nature and amount of guidance provided by the verbal instructions: *rule learning,* where students are given a sentence or question that verbally cues a solution to a problem, and *problem solving,* where the learner needs to discover the higher order rule without specific help. Our methodology fits into Gagné's definition of discovery learning because, rather than explicitly teaching students about the rules to solve each addition and subtraction case, our programs for Groups G and No-G allowed the students to manipulate the visual and symbolic representations with minimum or no guidance, respectively.

Discovery methods usually consist of presenting to students tasks that offer opportunities for constructing understandings such as having students interact with objects by exploring, manipulating, and testing hypotheses. The case for using discovery-based multimedia environments, where students are given the opportunity to construct understandings by interacting and manipulating the multiple representations, is based on the idea that discovery learning cultivates students' natural curiosity. It is consistent with Piaget's (1970) concept of the learner as naturally active, and more specifically, it is viewed as a powerful mechanism for cognitive change. Discovery-based multimedia allows students "opportunities to tinker with mathematical objects just as they might tinker with mechanical objects" (Cuoco & Goldenberg, 1996, p. 17). For

example, in our discovery-based multimedia game, students could test their thinking and immediately see the results by typing in an answer to a problem after manipulating the bunny's movements on the number line. Combining opportunities to manipulate the multiple representations on a large set of instances with computer feedback allows students to generalize and draw principles (Baxter, 1995; E. L. Ferguson & Hegarty, 1995; Thompson, 1992).

The growing consensus on the benefits of concrete manipulation (Berk, 1999), hands-on activities (Diem, 2001), and active learning (Cox, 1997; Mayer, 2001) lends support to the idea of developing discovery-based multimedia games such as the one used in the present study. However, Gagné (1965) warned about the potential inappropriate use of discovery learning where problem solving is treated "as though it could be achieved with a minimum of instructions and prerequisite knowledge of the rules" (p. 226). For instance, when students learn science in classrooms with pure discovery methods (where students make sense of the instructional materials without any help from the teacher), they often become lost and frustrated, and their confusion can lead to misconceptions (Brown & Campione, 1994; Hardiman, Pollatsek, & Weil, 1986). In addition, because false starts are common, unguided discovery may not be time efficient (Schauble, 1990).

In the area of multimedia learning, examples of discovery learning research include the study of Tarmizi and Sweller (1988), who found that students learned better when guidance in the form of subgoals or worked examples was provided during mathematics problem solving; the study of Mayer, Mautone, and Prothero (2002), who found that students learned more deeply from a computer-based geology game when they received explicit guidance concerning how to visualize relevant geological structures; and the work of Moreno (2004), who found that students learned better when explanatory rather than corrective feedback alone was provided in a multimedia game about environmental science. To our knowledge, there are no studies in the area of discovery-based multimedia that have examined the role of verbal guidance in helping students make sense of visual and symbolic representations. The present study set out to investigate this issue by comparing the effectiveness of two modified versions of discovery learning: with verbal guidance, where students were provided with an explanation about the relationship between the symbolic and visual representations of the arithmetic procedure, and without verbal guidance, where students needed to discover the relationship between the symbolic and visual representations of the arithmetic procedure by themselves.

Cognitive Theory of Multimedia Learning

Our research and program design were guided by a CTML consisting of the following main ideas (Mayer, 2001): (a) dual coding—in which the representation and processing of information concerning verbal and nonverbal materials are handled cognitively by separate subsystems (Clark & Paivio, 1991; Paivio, 1986), (b) dual processing—in which working memory includes independent auditory and visual working memories (Baddeley, 1992), (c) limited capacity—in which the processing capacities of learners are severely limited (Chandler & Sweller, 1991), and (d) active learning—in which meaningful learning occurs when learners select, organize, and build coherent connections of new information with prior knowledge (Mayer, 2001; Mayer & Moreno, 2003; Mayer &

Wittrock, 1996). On the basis of this model, Mayer and colleagues (Mayer, 2001; Mayer & Moreno, 2003; Moreno & Mayer, 2002) have identified a set of principles for the design of multimedia learning environments that are consistent with empirical research on multimedia learning and can be applied to new multimedia learning situations. We used these principles to make predictions about the potential benefits of having students learn with verbal guidance in a discovery-based multimedia game.

First, according to the *multimedia principle*, students learn better from words and pictures than from pictures or words alone (Mayer, 2001; Paivio, 1986). Because our guided version included verbal explanations in addition to the number sentence and number-line representations, CTML would have predicted that students in Group G should learn better than students in Group No-G, by encoding the additional verbal representation. This principle is also consistent with multiple-representation theory (Kozma et al., 1996; Moreno & Mayer, 1999b).

Second, because our guided version of the game included spoken explanations that were presented simultaneously with the visual materials and no irrelevant words or pictures were included, it was consistent with *modality, temporal contiguity,* and *coherence principles* of CTML. According to the modality principle, students learn better from words and pictures when words are presented in the auditory rather than visual modality because using independent processing channels to integrate the verbal and nonverbal materials prevents overloading the learner's limited visual working memory capacity (Baddeley, 1992). Thus, we expected our spoken explanations not to overload but rather to increase students' effective working memory capacity by using two rather than solely one processing channel (Moreno & Mayer, 1999a).

According to the temporal contiguity principle, when giving explanations, corresponding words and pictures should be presented simultaneously rather than successively so that auditory and visual information can be held in working memory at the same time and meaningful connections between the representations can be made (Mayer, 2001). With this principle in mind and to ensure that the proper connections between the verbal and nonverbal materials were made, we synchronized the spoken explanations with the corresponding movements of the bunny and the highlighting of the relevant portion of the number sentence in our guided version of the game. Finally, according to the coherence principle, students learn better when extraneous material is excluded from rather than included in a multimedia presentation (Chandler & Sweller, 1991; Moreno & Mayer, 2000). Our guided treatment was consistent with this principle because it included only words that corresponded to the relevant concepts that were represented visually in the multimedia game. In short, our guided computer game consisted of a narration, animation, and symbols that were temporally coordinated and concise.

Lastly, CTML's *individual-differences principle* states that design effects are different for learners of different cognitive capacities (Just & Carpenter, 1992). This last principle points out the limitations that working memory capacity may present to learning with discovery-based environments, where learners need to select, hold, and organize the different representations of knowledge before they can integrate them into their long-term memory (Mayer, 2001). According to Mayer (2001), even when the principles of multimedia design are applied, multiple-representation environments tend to help some kinds of learners (i.e., high prior

knowledge, high spatial ability) more than others (i.e., low prior knowledge, low spatial ability). In the remainder of this section, we introduce two additional potential sources of individual differences in learning from multiple representations and derive predictions based on CTML.

Computer Experience

According to CTML, there are some task-specific skills that derive from interacting with a multimedia program (such as clicking on buttons, typing in numbers, etc.) that may impose an extra cognitive load if students are not very familiar with computers and the tasks involved in the interaction have not become automatic yet (Chandler & Sweller, 1991). This was particularly relevant for the present study because many participants were children who came from low-income, non-English backgrounds with very few having access to computers at home. Consequently, we predicted that high-computer-experience students would learn more from our multimedia game than low-computer-experience students, especially with guidance.

Latency to Respond

Cognitive tempo refers to a person's tendency to reflect or not reflect before responding on a problem-solving task with response uncertainty (Jonassen & Grabowski, 1993; Kagan, 1966). The instrument most often used to measure cognitive tempo is the Matching Familiar Figures Test (MFFT; Kagan, Rosman, Day, Albert, & Phillips, 1964). In this test, time to the first response (latency) and number of overall errors are computed. We were particularly interested in examining the role that students' latency to respond may play when learning from multiple representations. We hypothesized that students with higher latencies to respond would be more likely to learn from our multimedia game by increasing the time to reflect on the relationships between the multiple representations, especially when guided during learning.

Method

Participants and Design

The participants were 61 students from three fifth- and sixth-grade classrooms of an elementary school in southern California who lacked substantial prior knowledge about addition and subtraction of signed num-

bers. All participating teachers had indicated that students had not been instructed on the topic yet. Thirty students participated in the no-G group (18 female, 12 male), and 31 students participated in the G group (20 female, 11 male).

Materials and Apparatus

For each participant, the paper-and-pencil materials consisted of a computer experience questionnaire, an arithmetic pretest and posttest, and the MFFT for children (Kagan et al., 1964). Computer experience was assessed by using a 10-item knowledge checklist. It consisted of instructions to "please answer the following questions and place a check mark next to the items that apply to you." The first 4 items were listed: "Do you have a computer at home? ____. If you do, how often do you use the computer at home?" "Do you use the computers at school? ____. If you do, how often do you use the computers at school?" "Do you go to the Boys and Girls Club to use the computers? ____. If you do, how often do you use the computers at the Boys and Girls Club?" "Do you go to the Teen Center to use the computers? ____. If you do, how often do you use the computers at the Teen Center?" Following each of these items, a space for a check mark next to the choices "Every day," "Twice a week," "Once a week," and "Less than once a week" was presented. The last 6 items of the checklist required students to place a check mark next to any of the following statements that applied to them: "I know what the World Wide Web is." "I navigated the World Wide Web before." "I know how to look for information in the World Wide Web." "I know how to play computer games." "I know what email is." "I sent an email message before."

The pretest and posttest were identical and consisted of an 8.5-in. × 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 numbers (P + P), addition of two negative 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 with three operands of the type N + P − N = and P − P + N = were also included. Table 1 lists the problems used in the pretest and posttest.

For each participant, the computer-based training materials consisted of a series of four interactive multimedia programs on addition and subtraction of signed integers and a computerized word-problem transfer test. All programs 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. Within each set, problem types were randomly ordered. The 16 problems contained in each of the

Table 1
Problems Used on Pretest/Posttest and in Each Learning Session

Items	A + B	−A + B	A + −B	A − B	−A + −B	−A − B	A − −B	−A − −B	3 operands
Pretest/posttest	3 + 1 =	−4 + 2 =	2 + −9 =	5 − 7 =	−5 + −4 =	−2 − 3 =	1 − −8 =	−5 − −6 =	−3 + 8 − −2 =
	7 + 2 =	−6 + 9 =	6 + −3 =	2 − 5 =	−8 + −1 =	−3 − 1 =	7 − −2 =	−8 − −6 =	4 − 7 + −5 =
Session 1	2 + 5 =	−3 + 5 =	4 + −5 =	5 − 7 =	−4 + −3 =	−5 − 1 =	4 − −5 =	−2 − −6 =	
	2 + 3 =	−3 + 7 =	5 + −6 =	4 − 5 =	−2 + −3 =	−4 − 2 =	1 − −8 =	−1 − −7 =	
Session 2	4 + 1 =	−1 + 4 =	3 + −4 =	3 − 2 =	−5 + −4 =	−1 − 3 =	2 − −6 =	−9 − −6 =	
	6 + 2 =	−4 + 2 =	7 + −5 =	6 − 3 =	−6 + −2 =	−2 − 3 =	4 − −3 =	−7 − −5 =	
Session 3	1 + 7 =	−5 + 1 =	8 + −7 =	6 − 4 =	−1 + −8 =	−7 − 2 =	3 − −5 =	−6 − −2 =	
	8 + 1 =	−9 + 1 =	1 + −4 =	7 − 2 =	−3 + −3 =	−3 − 2 =	2 − −3 =	−5 − −6 =	
Session 4	4 + 2 =	−7 + 2 =	2 + −9 =	1 − 4 =	−2 + −5 =	−2 − 2 =	1 − −3 =	−3 − −6 =	
	3 + 1 =	−2 + 4 =	7 + −2 =	6 − 3 =	−1 + −1 =	−6 − 2 =	5 − −2 =	−9 − −3 =	

four training sessions were identical for the no-G and G groups and are shown in Table 1. The computer program was designed to create a personal log in the computer's hard disk that recorded all of the students' interactions during the training session (e.g., the typed-in answers to the problems, the buttons clicked on the screen, the number of times students asked for a Spanish translation, and the exact time of each event). The programs were developed using Director 4.04 (Macromedia, 1995) and SoundEdit 16 (Macromedia, 1997).

The word-problem test contained a set of 16 word problems that were presented on the computer. Two word problems representing each of the eight problem types used in the training sessions were presented in random order. Students were instructed to click on the sentence that they considered best represented with words what the number sentence represented in symbols (among a set of four options). For example, for the problem "What could the following number sentence mean: $5 + 7 = \_\_$," students' options were "You lost 5 pounds yesterday and you lost 7 more pounds today," "You gained 5 pounds yesterday and you lost 7 pounds today," "You lost 5 pounds yesterday and you gained 7 pounds today," and "You gained 5 pounds yesterday and you gained 7 more pounds today." A box with hints was given to students on the lower right corner of the computer screen. For the above-described problem, the hints were "Positive numbers stand for weight you gain" and "Negative numbers stand for weight you lose." The hint box was designed as a scaffold to help students map the new abstract concept of a negative number to a concrete real world example. Spanish and English versions of the test were presented by on-screen text and narration. A button labeled *Español* on the lower right corner allowed students to access a Spanish translation of the word problem. Students' answers were recorded electronically by a computer-generated log. The computerized transfer test was developed using Director 4.04 (Macromedia, 1995). The apparatus consisted of 30 Sony headphones and Macintosh computer systems that included a 15-in. monitor and an average 40-megabyte internal hard disk. The computers were located in the elementary school's computer lab.

Procedure

For each one of the three participating classes, half of the students were randomly assigned to be in the no-G group and the other half to be in the G group. Both classes had identical mathematics curricula, and signed arithmetic was not covered in any of the classes. However, all participating students had prior understanding of the number line, as it was a common tool used in all classrooms to teach addition and subtraction of natural numbers. 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 the computer experience questionnaire and the paper-and-pencil pretest during regular class time. All tests were administered individually to all students on the same day. Students who scored at or below the median score (9 points) on the computer experience questionnaire were classified as low computer experience (15 in Group No-G and 16 in Group G), and students who scored above the median score were classified as high computer experience (15 in Group No-G and 15 in Group G).

Second, on the same day, each student was given the MFFT. One by one, students were called by the experimenter and taken to a quiet room in the school for its administration. The test consisted of a visual discrimination task with 12 items. On each trial, a black and white drawing of a familiar object was presented by the experimenter along with six similar pictures. Students were required to point to the picture that was identical to the original. For each item, latency to the first response was recorded using a stopwatch. If the first response was incorrect, the participant was asked to try again, and the total number of attempts until the correct match was identified was recorded. Because we were interested in examining how the latency component of cognitive tempo is associated with individual differences in learning from multiple representations, students scoring above the

median latency score of $-.16$ were classified as high latency (12 in Group No-G and 19 in Group G), and those scoring at or below the median latency score were classified as low latency (18 in Group No-G and 12 in Group G). A Latency × Computer Experience cross-tabulation showed that 17 students were classified as high computer experience/high latency, 13 were classified as high computer experience/low latency, 14 were classified as low computer experience/high latency, and 17 were classified as low computer experience/low latency.

Third, 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 lab, with each student seated at a Macintosh computer system. Both no-guidance and guidance versions of the program had been loaded on each computer's hard drive prior to the students' arrival. Students were randomly assigned a computer for every session, and the experimenter would start the corresponding version (i.e., no-guidance or guidance) before the student was seated. Next, students were instructed to put on the headphones and wait for the experimenter to start the program. Students were asked to solve all the problems in the multimedia game and, once all problems were solved, to review problems of their choice until the class time was over.

During each session, all students solved the 16 signed-arithmetic problems, working at their own rates as described above. After completing the four training sessions, all participants were given the paper-and-pencil posttest and the computerized word-problem transfer test and were asked to explain the solution to four example problems during an individual interview with the experimenter. All tests were administered individually, and the explanations given by students were audiotaped and later transcribed.

Scoring

A scorer, not aware of the treatment condition of each participant, determined the computer experience, pretest, and posttest scores. The computer experience score was calculated by adding the level of experience the participant checked on the self-assessment for the first four questions to the number of domain-related activities that the participant checked on the last six. A negative answer to any of the first four questions counted as 0. For positive answers, checking on *every day* counted as 4, *twice a week* as 3, *once a week* as 2, and *less than once a week* as 1. Although this procedure permits a maximum possible score of 22 points, the actual maximum possible was 20 points because students used computers at school only once a week. The minimum reported score was 3, and the maximum reported score was 15. The mean score and standard deviation for high- and low-computer-experience students were $Ms = 11.57$ and 6.26, $SDs = 1.25$ and 1.90, respectively.

For each student, we obtained an arithmetic pretest score by adding the number of correct answers produced on the pretest and a posttest score by adding the number of correct answers produced on the posttest. In addition, we recorded the following scores from the computer logs: first, the number of correct answers on each of the four training sessions resulting from adding all right answers that were typed by students during their first attempt to solve each problem; second, the number of times students in the G group clicked on the *Español* button during each training session; third, the number of right answers produced on the computerized word-problem test; and fourth, the number of times students clicked on the *Español* button during the computerized word-problem test. On the basis of the school records for their English language proficiency level, students were classified as English proficient (EP) or limited English proficient (LEP). LEP students were students who met two criteria. First, they were identified as coming from households where a non-English language was spoken based on a school-mandated survey administered to parents of enrolled students. Second, LEP students were found to score at the LEP level on the Language Assessment Scale—Oral English proficiency assessment (LAS; De Avila & Duncan, 1990). EP students were those coming from households where the English language was spoken, based on the school-

Table 2
Mean Scores and Corresponding Standard Deviations on Pretest, Posttest, Training Sessions, and Transfer for Two Groups

					Score										
							Session								
	Pretest		Posttest		1		2		3		4		Transfer	
Group	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
No-G	2.00	1.36	3.97	2.31	6.97	3.64	8.93	3.14	10.00	3.23	9.80	2.93	7.43	2.73
G	2.03	1.22	5.13	2.51	8.52	3.96	10.39	2.82	10.45	3.04	10.39	3.63	8.84	3.31

Note. Potential scores ranged from 0 to 18 for pretest and posttest, from 0 to 16 for sessions, and from 0 to 16 for transfer. No-G = no guidance; G = guidance.

mandated parent survey, plus those who came from households where a non-English language was spoken but who did not score at the LEP level on the LAS. Fourteen EP and 16 LEP students served in Group No-G, and 15 EP and 16 LEP students served in Group G.

Finally, the think-aloud protocols from each participant were scored by a graduate student who was unaware of the group to which each student had been assigned in the experiment. Roxana Moreno served as a second rater of a random sample consisting of 50% of students' protocols. The kappa coefficient used to estimate interrater reliability (G. A. Ferguson, 1981) from the sample was determined to be .94. Differences between raters were solved by agreement.

Protocols

The problems used during the think-aloud process consisted of one example for each of the four most difficult problem types (P + N, P − N, N − P, and N + P). We computed the pretest score for each problem type and chose the four categories where the average score was the lowest. During the protocol session, the experimenter asked students to explain in their own words (either in English or Spanish) how they would solve each problem. Each problem was represented on a sheet of paper with the symbolic sentence and a number line identical to the number line used in the training sessions. Students were told that they could move along the number line to help them solve the problem if they needed to do so. Students' protocols were transcribed and scored. The scoring consisted of comparing students' explanations for solving each problem with the steps that had been explicitly provided by the guidance program. For example, for the problem "3 − −6 = __," the steps were, first, find the starting point—three means go to three; second, find the operation—to subtract means to face left; third, find the direction to jump—negative six means jump back six steps. Therefore, as there were three steps involved in solving the problem, the protocol score ranged from 0 to 3 for each problem.

Results

First, a two-way contingency table analysis was conducted to examine the independence between computer experience and latency. The variables were not found to not be significantly related, Pearson $\chi^2(1, N = 61) = .81$, $p = .37$. Next, for each individual-difference measure (computer experience and latency), the data were subjected to (a) a two-way analysis of variance (ANOVA) with treatment group and the individual difference as the between-subjects factors and pretest and posttest scores as the within-subject factor, (b) a two-way ANOVA with treatment group and the individual difference as the between-subjects factors and the scores obtained in each training session as the within-subject

factor, and (c) a two-way ANOVA with the between-subjects factors being treatment group and the respective individual-difference measure and transfer score as the dependent measure. All significant interactions were followed up with supplemental Tukey tests (with alpha at .05), and effect sizes were assessed by η^2. Table 2 summarizes the mean scores and corresponding standard deviations on pretest, posttest, training sessions, and transfer for each treatment group.

Pretest and Posttest Scores

The first four columns of Table 3 summarize the mean pretest and posttest scores and standard deviations for different computer-experience and latency groups in each treatment condition.[1] Using group and computer experience as between-subjects factors, the repeated-measure ANOVA revealed a main pretest-to-posttest effect, $F(1, 57) = 119.79$, $MSE = 202.21$, $p < .0001$. Students' posttest scores were significantly higher than their pretest scores ($Ms = 4.66$ and 2.06, $SDs = 2.33$ and 1.28, respectively). The effect size was .68. In addition, we found a significant interaction between pretest and posttest and group, $F(1, 57) = 6.30$, $MSE = 10.63$, $p = .02$. Consistent with our predictions, although both groups showed learning of the arithmetic procedure over time, the posttest scores of Group G were significantly larger than those of Group No-G (the effect size was .10). Finally, there was a significant interaction between group and computer experience, $F(1, 57) = 5.44$, $MSE = 28.37$, $p = .02$, indicating that high-computer-experience students in Group G scored significantly higher than the rest of the subgroups, which did not differ significantly from each other. The effect size was .09. No main effect for treatment group or computer experience and no other significant interactions were found.

Similar to the case of computer experience, using group and latency as between-subjects factors, the repeated-measure ANOVA revealed a main pretest-to-posttest effect, $F(1, 57) = 105.26$, $MSE = 179.74$, $p < .0001$, and a significant interaction

---

[1] Although students were trained in addition and subtraction of integers with two operands, the pretest and posttest included 2 three-operand items (see Table 1). A separate ANOVA was conducted on these two items alone, to find a "far transfer effect." However, because no significant differences between groups, individual measures, or interactions were found, we collapsed all 18 items in the pretest and posttest analyses.

Table 3
Mean Scores and Corresponding Standard Deviations on Pretest, Posttest, Training Sessions, and Transfer for Different Computer Experience and Latency Groups

	Score													
					Session									
	Pretest		Posttest		1		2		3		4		Transfer	
Group & classification	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
No-G														
Low CE	2.00	1.22	4.31	1.89	6.92	3.45	8.00	2.92	9.46	3.31	9.84	3.16	7.53	2.22
High CE	2.00	1.50	3.71	2.62	7.00	3.89	9.65	3.20	10.41	3.20	9.76	2.84	7.35	3.12
Low latency	1.89	1.32	3.67	2.50	7.22	3.35	8.94	3.19	9.22	3.14	9.11	2.95	7.22	2.69
High latency	2.17	1.47	4.42	2.02	6.58	4.17	8.92	3.20	11.17	3.13	10.83	2.69	7.75	2.86
G														
Low CE	1.61	1.03	4.17	2.04	6.67	3.97	9.06	2.48	9.17	2.68	8.61	2.68	7.17	2.92
High CE	2.62	1.26	6.46	2.57	11.08	2.14	12.23	2.20	12.23	2.65	12.85	3.39	11.15	2.30
Low latency	1.58	1.16	3.92	1.44	6.75	3.72	9.08	2.23	9.50	3.09	8.42	2.54	6.67	2.67
High latency	2.32	1.20	5.90	2.77	9.63	3.77	11.21	2.89	11.05	2.93	11.63	3.71	10.21	2.95

Note. Potential scores ranged from 0 to 18 for pretest and posttest, from 0 to 16 for sessions, and from 0 to 16 for transfer. No-G = no guidance; G = guidance; CE = computer experience.

between pretest and posttest and group, $F(1, 57) = 3.78$, $MSE = 6.46$, $p = .05$. In addition, there was a significant latency effect, $F(1, 57) = 4.69$, $MSE = 25.42$, $p < .05$, indicating that high-latency students had higher mean scores than low-latency students ($Ms = 3.79$ and 2.76, $SDs = 1.72$ and 1.56, respectively). The effect size was .08. No other significant interactions were found.

Training Session Scores

The middle eight columns of Table 3 summarize the mean session scores and standard deviations for different computer-experience and latency groups in each treatment condition. Using group and computer experience as between-subjects factors and training session as a within-subject factor, we found a main effect for session, $F(3, 171) = 17.11$, $MSE = 88.22$, $p < .0001$. The effect size was .34. Students' mean score for Session 1 was significantly lower than the mean scores for Sessions 2, 3, and 4, which did not differ from each other ($Ms = 7.75, 9.67, 10.23,$ and 10.10, $SDs = 3.85, 3.05, 3.12,$ and 3.29, for Sessions 1 to 4, respectively). In addition, there was a main group effect, $F(1, 57) = 4.49$, $MSE = 109.33$, $p < .05$, with students in Group G having session scores higher than students in Group No-G ($Ms = 9.94$ and 8.93, $SDs = 2.88$ and 2.67, respectively; effect size of .07), and a main effect for computer experience, $F(1, 57) = 11.66$, $MSE = 284.70$, $p = .001$, indicating that high-experience students had higher mean session scores than low-experience students ($Ms = 10.46$ and 8.45, $SDs = 2.97$ and 2.26, respectively; effect size of .17). Finally, there was a significant interaction between group and computer experience, $F(1, 57) = 5.77$, $MSE = 140.82$, $p = .02$. High-computer-experience students in Group G had significantly higher overall session means than the rest of the groups, which did not differ significantly from each other. The effect size was .09. There were no other significant interactions.

Using group and latency to respond as between-subjects factors and training session as a within-subject factor, we found a main effect for session, $F(3, 171) = 20.30$, $MSE = 101.86$, $p < .0001$,

and a main effect for latency, $F(1, 57) = 5.19$, $MSE = 148.47$, $p < .05$. High-latency students had higher mean session scores than low-latency students ($Ms = 10.30$ and 8.55, $SDs = 2.78$ and 2.57, respectively; effect size of .08). There were no other significant interactions.

Transfer to Word Problems

The last two columns of Table 3 summarize the mean transfer scores and standard deviations for different computer-experience and latency groups in each treatment condition. Using computer experience and group as between-subject factors, the ANOVA revealed a group effect in favor of students in Group G, $F(1, 57) = 5.90$, $MSE = 43.84$, $p < .05$ (effect size of .09), and a main computer-experience effect, $F(1, 57) = 7.26$, $MSE = 53.88$, $p = .009$ (effect size of .11). High-computer-experience students were better able to solve word problems as compared with low-computer-experience students ($Ms = 9.00$ and 7.32, $SDs = 3.35$ and 2.61, respectively). We also found a significant interaction between computer experience and group, $F(1, 57) = 8.74$, $MSE = 64.91$, $p = .005$ (effect size of .16), indicating that, similar to the case of pretest to posttest and training session improvements, the transfer scores of high-computer-experience students in Group G were significantly higher than the transfer scores of the rest of the groups.

Using latency and group as between-subject factors, the ANOVA revealed a latency main effect, $F(1, 57) = 7.65$, $MSE = 60.32$, $p = .008$ (effect size of .12). High-latency students were better able to solve word problems as compared with low-latency students ($Ms = 9.25$ and 7.00, $SDs = 3.12$ and 2.65, respectively). In addition, a significant interaction between latency and group was found, $F(1, 57) = 4.20$, $MSE = 33.10$, $p < .05$. The transfer scores of high-latency students in Groups No-G and G did not differ from each other and were higher than the transfer scores of low-latency students in Groups No-G and G, and low-latency

students in Group G had higher transfer scores than low-latency students in Group No-G. The effect size was .07.

The Role of Verbal Guidance

So far, we have reported the findings related to the questions "Do students learn more with guided versus pure discovery multimedia? And are there important individual differences in which discovery-based multimedia games help some kinds of learners more than others?" However, still open is the question "What is the role of verbal guidance in promoting learning from a discovery-based multimedia game?" In the remainder of this section, we report the data from student think-aloud protocols and first language use to help answer this question.

Student think-aloud protocols. According to CTML, a possible way that guidance helps learning is by encoding the procedure as an additional verbal representation. If this were the case, we should have found evidence that students who learned with guidance used the verbal explanations presented during training more often than those not provided with the verbal explanations while solving problems. To test this hypothesis, we recorded the mean number of times that students used the guiding steps provided to the G group to solve four problems during the protocol session. Using the mean number of steps reported during problem solving on the four problems as a dependent measure, a one-factor ANOVA failed to reveal a main treatment effect. Students in Group G did not use the guiding steps more than students in Group No-G while solving problems. The mean number of verbal steps reported by both groups was very low (Ms $= 4.00$ and 2.95, SDs $= 3.35$ and 3.55, for G and no-G groups, respectively), with only one student from each group achieving the maximum potential score of 12. Most students in both groups produced alternative explanations for solving the arithmetic problems.

Students' choice of language. As part of our interest in the role of verbal explanations in discovery-based multimedia, we examined whether students who learned with guidance would choose to learn the arithmetic procedure in their first or second language given the option to do so. The idea was that non-English background learners with limited proficiency in English would find it easier to decode and understand verbal problem information in their first language rather than English (Durán, 1985). For example, bilinguals respond more slowly in a less familiar language than in their native language to simple oral instructions (Dornic, 1977), incur greater communicative workload when the medium of communication is in their nondominant language (Dornic, 1980;

Hamers & Lambert, 1972; Preston & Lambert, 1969), and read sentence-length materials slower in a less familiar versus a more familiar language (Kolers, 1966; Macnamara, 1967). In examining this issue, we built on research (Durán, 1985; Mestre, 1988; Solano-Flores & Trumbull, 2003) establishing the importance of understanding better how language background affects problem solving. Because students in Group G were given the option to listen to the explanations in English or Spanish, we expected that students would choose to listen to their dominant language during learning, with LEP learners showing a greater use of Spanish explanations over sessions as compared with EP learners. Accordingly, for each session and student, we recorded the number of times the students listened to Spanish explanations. Table 4 summarizes the mean use of Spanish explanations and corresponding standard deviations for LEP students and EP students on the four training sessions.

An ANOVA using English language proficiency (EP vs. LEP status) as a between-subjects factor, session number as a within-subject factor, and number of Spanish explanations requested as a dependent variable revealed a significant effect for English proficiency and for number of Spanish explanations, as well as a significant interaction, $F(1, 87) = 5.69$, $MSE = 122.58$, $p < .05$; $F(3, 87) = 9.25$, $MSE = 51.77$, $p = .0001$; and $F(3, 87) = 5.82$, $MSE = 32.56$, $p = .001$, respectively. LEP students used significantly more Spanish translations than EP students, and the number of Spanish explanations used by both groups in Sessions 1 and 2 was significantly higher than those used in Sessions 3 and 4. In addition, the number of Spanish explanations used by LEP students during Session 1 was higher than the number used during Sessions 2, 3, and 4, and the number of Spanish explanations used during Session 2 was higher than the number used during Sessions 3 and 4, which did not differ from each other. Finally, LEP students used more Spanish explanations during Sessions 1 and 2 than EP students during Sessions 1, 2, 3, and 4. The numbers of Spanish explanations per session used by EP students were very low and not significantly different from each other.

Discussion

Computers can be powerful aids to mathematics teaching and learning and are considered an important force behind the evolution of mathematics education (Heid, 1997; Wiest, 2001). When properly designed, multimedia games have the potential of facilitating mathematics learning and therefore can supplement other instructional methods (National Council of Teachers of Mathemat-

Table 4

Mean Use of Spanish Explanations and Corresponding Standard Deviations for EP and LEP Students on Four Training Sessions

| | Number of Spanish explanations | | | | | | | | | |
| | 1 | | 2 | | 3 | | 4 | | Overall | |
Group	M	SD	M	SD	M	SD	M	SD	M	SD
EP	0.40	0.83	1.27	2.55	0.00	0.00	0.00	0.00	0.42	1.41
LEP	5.38	6.31	2.94	4.52	0.81	2.01	0.50	1.75	2.41	4.46

Note. EP = English proficient; LEP = limited English proficient.

ics, 2000). For example, discovery-based multimedia math games can provide a medium to access and process information, model and explore mathematics, and conduct mathematical investigations at a student's own pace (Wiest, 2001). However, learning a new arithmetic procedure with multiple representations may be overwhelming to the novice learner (Moreno, 2004), especially when asked to select, organize, and integrate the multiple representations with prior knowledge and without guidance (Sweller, 1994). Does verbal guidance help students learn mathematics from a multiple-representation multimedia game? Are there individual differences for which a multimedia game helps some kinds of learners more than others? Two contributions of the present study are, first, to show that verbal guidance is an important method to promote learning with discovery-based multimedia consisting of multiple representations and, second, to point out that students' lower computer proficiency and latency to respond can undercut the potential benefits of learning with multimedia environments. Importantly, our findings extend the growing body of research showing that students learn more deeply from guided discovery than pure discovery methods to the area of multimedia learning (Brown & Campione, 1994; Hardiman et al., 1986; Schauble, 1990; Solter & Mayer, 1978).

Theoretical and Practical Implications

Theoretically, this research supports CTML in two ways. First, we have shown that although both discovery-based groups showed learning of the arithmetic procedure over time, the posttest scores of students in Group G were significantly larger than those of Group No-G. We interpret this interaction to support a CTML according to which providing verbal guidance in addition to visual and symbolic representations helps students understand the complex arithmetic procedure better than having students discover the relationship between the multiple representations on their own (Mayer, 2001; Moreno & Mayer, 1999b). Additional evidence supporting this interpretation was found by looking at the number of nonlearners—students who had zero or negative pretest to posttest gains. On the basis of this definition, we found 10 nonlearners, of whom 9 had trained in Group No-G and only 1 had trained in Group G. That is, 90% of students for whom training was ineffective had received no guidance.

The second way that the present study contributes to CTML is by showing that a discovery-based computer game may have positive learning effects only when students possess the necessary cognitive resources to process simultaneously in working memory the multiple representations (Chandler & Sweller, 1991; Sweller, 1994). More specifically, our work extends the individual-differences principle derived from CTML (Mayer, 2001) to two previously unexplored variables: computer experience and latency to respond. For multimedia learning to be effective, it is important to design the materials in a manner that maximizes students' opportunities to actively select, organize, and integrate multiple representations of knowledge. However, our findings show that even well-designed, principle-based multimedia games do not help all students learn equally well: High-computer-experience students (especially those learning with verbal guidance) and high-latency students outperformed their counterparts on session and transfer scores.

On the practical side, the most direct implication of this study is to show that discovery-based environments that use multiple representations should be designed to include verbal guidance in students' first language if they are to be used as an instructional tool to foster mathematical understanding. In addition, our individual-differences results suggest that in the teaching of mathematics with interactive multiple-representation games, the ideal learning situation might be, first, to bring students to a high level of computer proficiency and, second, to emphasize the importance of taking the time to reflect on a problem's solution before attempting an answer. Similar to past research, our results suggest that learners with low latencies to respond are at a seeming disadvantage in problem-solving situations (Readence & Bean, 1978a). However, instruction to delay responding, reflective modeling, or self-guidance techniques can be embedded in the computer environment to improve their performance (Readence & Bean, 1978b; Waring, Farthing, & Kidder-Ashley, 1999).

Limitations and Future Directions

It is important to note that, for both treatment conditions, students did not improve significantly after the second training session and that the session and posttest scores were very low as compared with the maximum potential score. One possible interpretation for this finding is that students may have mastered only a subset of integer problems after the second session of training and that more training sessions are needed to understand the procedure for all addition and subtraction cases. The lack of evidence of the use of the verbal steps in the four most difficult problem types during students' think-aloud protocols supports this idea further. Because the topic requires significant accommodation of the arithmetic procedure for natural numbers to the one for signed numbers, our results suggest that more time to reflect on the conceptual and procedural differences that exist between positive and negative numbers seems to be an important condition affecting learning more fully from our multimedia multi-representational environment. An alternative explanation is that discovery learning may not be as effective as other methods—such as direct instruction, when learners need to combine previously learned rules into an entirely new class of problems that contradict their prior knowledge. As Gagné stated, "The evidence of experimental studies concerning the use of discovery in problem solving certainly does not demonstrate that higher-order rules *must* be learned by discovery" (Gagné, 1965, p. 225). These hypotheses need to be examined in further research.

Finally, our research is limited because it has dealt with only one knowledge domain (i.e., addition and subtraction of integers), one kind of guidance (i.e., verbal explanations), and one kind of learner (i.e., elementary students). More research is needed to determine how to incorporate structured guidance using other domains, methods, and populations. Furthermore, it is possible that some of these results do not generalize to materials that are less challenging. When the material to be learned is complex, an unguided discovery method may be detrimental to learning by exhausting students' working memory capacity. Had the materials been easier, our results may have been different or even opposite. In this case, a discovery method with little or no verbal guidance may have provided effective conditions to foster students' learning. Lastly, a limitation of our study is that, of the wide realm of individual

differences in learning, we chose to focus on computer experience and latency to respond. Future research should extend the present study to other individual differences in learning from multiple-representation multimedia games.

References

Baddeley, A. (1992, January 31). Working memory. *Science, 255,* 556–559.

Baxter, G. P. (1995). Using computer simulations to assess hands-on science learning. *Journal of Science Education and Technology, 4,* 21–27.

Berk, E. G. (1999). Hands-on science: Using manipulatives in the classroom. *Principal, 78,* 52–55, 57.

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.

Brown, A., & Campione, J. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229–270). Cambridge, MA: MIT Press.

Bruner, J. S. (1961). The art of discovery. *Harvard Educational Review, 31*(1), 21–32.

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, Serial No. 246).

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.

Cox, B. D. (1997). The rediscovery of the active learner in adaptive contexts: A developmental-historical analysis of transfer of training. *Educational Psychologist, 32,* 41–45.

Cuoco, A. A., & Goldenberg, E. P. (1996). A role for technology in mathematics education. *Journal of Education, 178*(2), 14–32.

De Avila, E. A., & Duncan, S. E. (1990). *LAS, Language Assessment Scales, Oral administration manual, English, Forms 2C and 2D.* Monterey, CA: CTB/MacMillan/McGraw-Hill.

Diem, K. G. (2001, April). Turn any science topic into hands-on fun. *Science Scope, 15,* 46–49.

Dornic, S. (1977). Selective attention bilingual tasks. In S. Dornic (Ed.), *Attention and performance VI* (pp. 669–678). Hillsdale, NJ: Erlbaum.

Dornic, S. (1980). Language dominance, spare capacity, and perceived effort in bilinguals. *Ergonomics, 23,* 369–377.

Durán, R. P. (1985). Influences on language skills on bilinguals' problem solving. In S. Chipman, J. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Research and open questions* (pp. 187–207). Hillsdale, NJ: Erlbaum.

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

Ferguson, E. L., & Hegarty, M. (1995). Learning with real machines or diagrams: Application of knowledge to real-world problems. *Cognition and Instruction, 13,* 129–160.

Ferguson, G. A. (1981). *Statistical analysis in psychology and education.* New York: McGraw-Hill.

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: Erlbaum.

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.

Gagné, R. M. (1965). *The conditions of learning.* New York: Holt.

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.

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

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

Hamers, J. F., & Lambert, W. E. (1972). Bilingual interdependencies in auditory perception. *Journal of Verbal Learning and Verbal Behavior, 11,* 303–310.

Hardiman, P., Pollatsek, A., & Weil, A. (1986). Learning to understand the balance beam. *Cognition and Instruction, 3,* 1–30.

Heid, M. K. (1997). The technological revolution and the reform of school mathematics. *American Journal of Education, 106,* 5–61.

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.

Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences.* Hillsdale, NJ: Erlbaum.

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

Kagan, J. (1966). Developmental studies in reflection and analysis. In A. H. Kidd & J. L. Rivoire (Eds.), *Perceptual development in children* (pp. 487–522). New York: International University Press.

Kagan, J., Rosman, B. L., Day, K., Albert, J., & Phillips, W. (1964). Information processing in the child: Significance of analytic and reflective attitudes. *Psychological Monographs, 78*(1, Whole No. 578).

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

Kolers, P. A. (1966). Reading and talking bilingually. *American Journal of Psychology, 79,* 357–376.

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: Erlbaum.

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: Erlbaum.

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

Macnamara, J. (1967). The effects of instruction in a weaker language. *Journal of Social Issues, 23,* 121–135.

Macromedia. (1995). Director 4.04 [Computer program]. San Francisco: Author.

Macromedia. (1997). SoundEdit 16 [Computer program]. San Francisco: Author.

Mayer, R. E. (2001). *Multimedia learning.* New York: Cambridge University Press.

Mayer, R. E., Mautone, P. D., & Prothero, W. (2002). Pictorial aids for learning by doing in a multimedia geology simulation game. *Journal of Educational Psychology, 94,* 171–185.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38,* 43–52.

Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In R. Calfee & R. Berliner (Eds.), *Handbook of educational psychology* (pp. 47–62). New York: Macmillan.

Mestre, J. P. (1988). The role of language comprehension in mathematics and problem solving. In Rodney R. Cocking & Jose P. Mestre (Eds.), *Linguistic and cultural influences on learning mathematics: The psychology of education and instruction* (pp. 201–220). Hillsdale, NJ: Erlbaum.

Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback on discovery-based multimedia. *Instructional Science, 32,* 99–113.

Moreno, R., & Mayer, R. E. (1999a). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91,* 358–368.

Moreno, R., & Mayer, R. E. (1999b). Multimedia-supported metaphors for meaning making in mathematics. *Cognition and Instruction, 17,* 215–248.

Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology, 92,* 117–125.

Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology, 94,* 598–610.

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

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

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

Piaget, J. (1970). *The science of education and the psychology of the child.* New York: Orion Press.

Preston, M. S., & Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop color word task. *Journal of Verbal Learning and Verbal Behavior, 8,* 295–301.

Readance, J. E., & Bean, T. W. (1978a). Impulsivity-reflectivity and learning: An individual difference that matters. *College Student Journal, 11,* 367–374.

Readance, J. E., & Bean, T. W. (1978b). Modification of impulsive cognitive style: A survey of the literature. *Psychological Reports, 43,* 327–337.

Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology, 49,* 31–57.

Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher, 32,* 123–145.

Solter, A., & Mayer, R. E. (1978). Broader transfer due to guided discovery of number concepts in preschool children. *Journal of Educational Psychology, 70,* 363–37l.

Spiro, R., Feltovich, P., Jacobson, M., & Coulson, R. (1992). Knowledge representation, content specification, and the development of skill in situation-specific knowledge assembly: Some constructivist issues as they relate to cognitive flexibility theory and hypertext. In T. Duffy & D. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation* (pp. 121–127). Hillsdale, NJ: Erlbaum.

Sternberg, R. J., & Frensch, P. A. (1993). Mechanisms of transfer. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 23–38). Norwood, NJ: Ablex Publishing.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4,* 295–312.

Tarmizi, R. A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80,* 424–436.

Thompson, P. W. (1992). Notations, conventions, and constraints: Contributions to effective uses of concrete materials in elementary mathematics. *Journal for Research in Mathematics Education, 23,* 123–147.

Wagner, S., & Kieran, C. (Eds.). (1989). *Research issues in the learning and teaching of algebra.* Reston, VA: National Council of Teachers of Mathematics.

Waring, D. A., Farthing, C. B., & Kidder-Ashley, P. (1999). Impulsive response style affects computer-administered multiple-choice test performance. *Journal of Instructional Psychology, 26,* 121–128.

Wiest, L. R. (2001). The role of computers in mathematics teaching and learning. *Computers in the Schools, 17,* 41–55.

Wittrock, M. C. (1966). The learning by discovery hypothesis. In L. S. Shulman & E. R. Keislar (Eds.), *Learning by discovery: A critical appraisal* (pp. 33–75). Chicago: Rand McNally.