

# Multimedia Learning: Are We Asking the Right Questions?

Richard E. Mayer

*University of California, Santa Barbara*

How can we help students to understand scientific explanations of cause-and-effect systems, such as how a pump works, how the human respiratory system works, or how lightning storms develop? One promising approach involves multimedia presentation of explanations in visual and verbal formats, such as presenting computer-generated animations synchronized with computer-generated narration or presenting illustrations next to corresponding text. In a review of eight studies concerning whether multimedia instruction is effective, there was consistent evidence for a multimedia effect: Students who received coordinated presentation of explanations in verbal and visual format (multiple representation group) generated a median of over 75% more creative solutions on problem-solving transfer tests than did students who received verbal explanations alone (single representation group). In a review of 10 studies concerning when multimedia instruction is effective, there was consistent evidence for a contiguity effect: Students generated a median of over 50% more creative solutions to transfer problems when verbal and visual explanations were coordinated (integrated group) than when they were not coordinated (separated group). Finally, in a review of six studies concerning for whom multimedia instruction is effective, Attribute  $\times$  Treatment interactions indicated that multimedia and contiguity effects were strongest for low prior knowledge and high spatial ability students. Results are consistent with a generative theory of multimedia learning in which learners actively select, organize, and integrate verbal and visual information.

For the past 7 years, my colleagues and I at the University of California, Santa Barbara, have been examining methods for improving students' understanding of scientific explanations. Here's the kind of situation that got us interested. I ask a student who lacks experience in meteorology to read a short text lesson explaining how lightning storms develop. The passage contains approximately 600 words and includes a cause-and-effect explanation of the events leading to the formation of lightning. A few minutes later, I ask the student to answer a transfer question, that is, a question that requires reasoning about the information in the lesson. In spite of an earnest attempt to learn, he or she is unable to produce an acceptable answer and, in fact, performs no better than an equivalent student who has not received any training.

Why is it that a student can read or listen to every word of a scientific passage, including a cause-and-effect explanation, and yet not be able to use that information to solve problems? Our research has produced convincing evidence that presenting a verbal explanation of how a system works does not insure that students will understand the explanation. In our search for ways to help students understand scientific explanations, we have come to rely increasingly on what has been

called *multimedia learning*, through presenting explanations visually as well as verbally. Multimedia learning occurs when students receive information presented in more than one mode, such as in pictures and words. In recent years, the once near monopoly of verbally based modes of instruction has given way to the hypothesis that meaningful learning occurs when learners construct and coordinate multiple representations of the same material, including visual and verbal representations (Houghton & Willows, 1987; Mandl & Levin, 1989; Schnotz, 1993a; Schnotz & Kulhavy, 1994; Willows & Houghton, 1987). For example, Schnotz and Kulhavy (1994, p. vi) recently observed: "Many studies have shown that graphics can make communication and learning more effective, but we have only recently begun to understand better why and under what conditions they are really effective."

In defining multimedia learning it is useful to distinguish among *delivery media*, *presentation modes*, and *sensory modalities*. Delivery media refers to the system used to present instruction, such as a book-based medium versus a computer-based medium. This distinction is examined in the section on Question 1. Presentation modes refer to the format used to represent the presented instruction, such as words versus pictures. This distinction is examined in the sections on Questions 2, 3, and 4. Sensory modality refers to the information processing channel that a learner uses to process the

information, such as acoustic versus a visual information processing. This distinction is examined in the section on Question 5. For example, a description of how something works can be delivered via printed text in a book or printed text on a computer screen (i.e., two different media), in the form of a series of illustrations or in the form of a series of printed statements (i.e., two different modes), or as printed words or spoken words (i.e., two different modalities). Although early research on media effects focused on the first distinction, I focus mainly on the second distinction in my research. Thus, it may be more accurate to use the term "multimodal learning" than "multimedia learning."

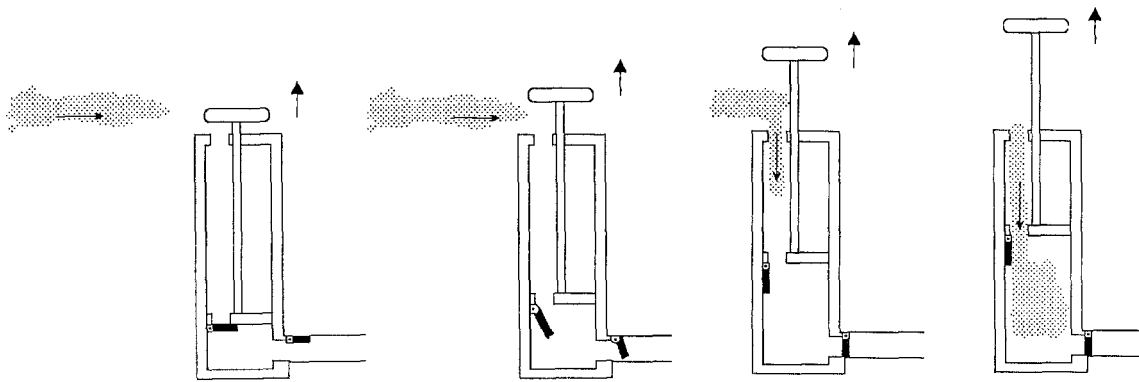
In this article, I begin with an example of a multimedia learning scenario. Next, I describe what I mean by scientific explanation and I present a theory of how students understand a scientific explanation. Finally, I review research on multimedia learning of scientific explanations, focusing on a series of 10 experiments that test the conditions under which the presentation of visual and verbal explanations is most effective in promoting student understanding. Because several

excellent reviews of media research already exist (Clark & Salomon, 1986; Kozma, 1991; Ross, 1994; Wetzel, Radtke, & Stern, 1994), I have focused this article on an integrative review of multimedia research conducted in my laboratory at Santa Barbara.

### A MULTIMEDIA LEARNING SCENARIO

Consider the following multimedia learning scenario in a computer-based environment: You are sitting in front of a computer screen, and by using the mouse you ask the computer to explain how a bicycle tire pump works. On the screen you see a short, simple animation depicting the operation of a bicycle tire pump, and at the same time you hear a corresponding narration of the actions involved in the operation of the pump as they are shown. Figure 1 presents some frames from a tire pump animation along with the complete sound tract. This scenario is representative of the kind of animation materials currently available in electronic encyclopedias such

"When the handle is pulled up, the piston moves up, the inlet valve opens, the outlet valve closes, and air enters the lower part of the cylinder."



"When the handle is pushed down, the piston moves down, the inlet valve closes, the outlet valve opens, and air moves out through the hose."

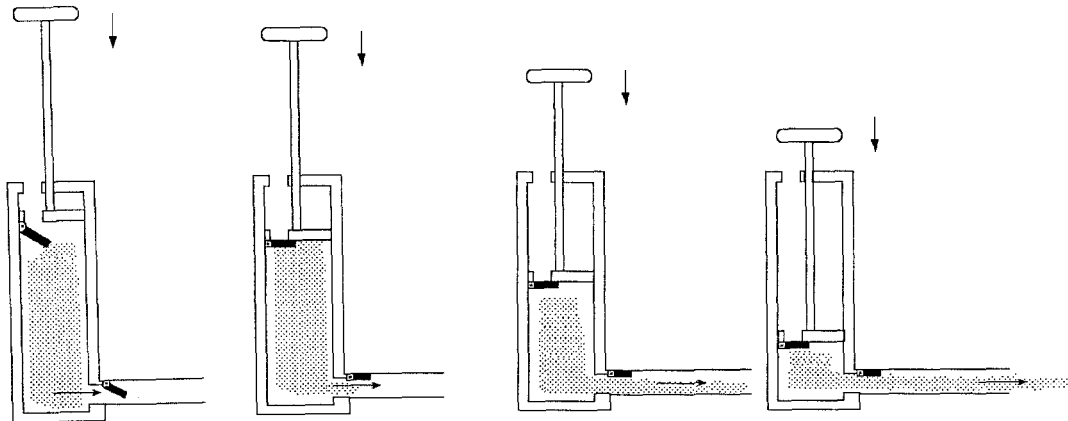


FIGURE 1. Selected animation frames and corresponding narration from a multimedia lesson on how a tire pump works. (Adapted from *The World Book Encyclopedia*, Vol. 15, p. 904, 1994. Chicago: World Book, Inc. Copyright 1994 by The World Book Encyclopedia. By permission of publisher.)

as *Compton's Interactive Encyclopedia*, *Grolier Multimedia Encyclopedia*, and *Microsoft Encarta*.

A corresponding example of a multimedia learning scenario occurs in a textbook-based environment: You look up the term "bicycle tire pump" in an encyclopedia and find a passage that describes how a bicycle tire pump works. On the same page is an illustration containing several frames that depict the state of the parts in the pump when the handle is pushed down and pulled up. Figure 2 presents an excerpt of the tire pump text and illustrations. This scenario is representative of the kind of illustration materials currently available in conventional encyclopedias such as the *World Book Encyclopedia*.

These are examples of multimedia learning experiences because the learner receives verbal information (such as printed text or narration) and visual information (such as illustrations or animation). We can assess the cognitive consequences of multimedia learning by asking the learner to remember the presented information (i.e., retention test) or to answer questions that require reasoning with the presented information (i.e., transfer test). Examples of transfer test items are listed in Table 1.

Because understanding of scientific explanations is the main focus of our research, I focus on problem-solving transfer as the major dependent measure of interest. The ability to transfer learning to the solution of new problems has long been recognized as a better measure of learner understanding than mere retention of the material (Wertheimer, 1959). In our studies we score each student's problem-solving answers by tallying the number of acceptable solutions on each question, that is, solutions that are both correct and creative. For example, an acceptable answer for the question about reliability is to use airtight seals or to use a backup system, whereas an unacceptable answer is to use better materials; an acceptable

TABLE 1  
Some Problem Solving Transfer Questions

What can be done to make a pump more reliable, that is, to make sure it would not fail?
What could be done to make a pump more effective, that is, to move more liquid or gas more rapidly?
Suppose you push down and pull up the handle of a tire pump several times but no air comes out. What could have gone wrong?
Why does air enter a tire pump? Why does air exit from a tire pump?

answer to the question about effectiveness is to use a larger cylinder or to press down harder, whereas an unacceptable answer is be sure to follow directions; an acceptable answer to the question about troubleshooting is that there may be a hole in the cylinder or a stuck valve, whereas an unacceptable answer is that the pump is broken; and an acceptable answer for the question about why air enters and exits is that a vacuum accounts for air entering or compression accounts for air exiting the cylinder whereas an unacceptable answer is that air is pushed out. A student's problem-solving score is the total number of acceptable solutions generated across the set of problems.

## THE NATURE OF SCIENTIFIC EXPLANATION

An important first step in designing effective instruction concerning scientific explanations is to have a clear idea of the nature of scientific explanation. Mayer (1992b) identified three views of scientific explanation: description of phenomena, induction of rules, and invention of models. According to the description of phenomena view, a scientific explanation exists when a scientist observes events and describes them. In

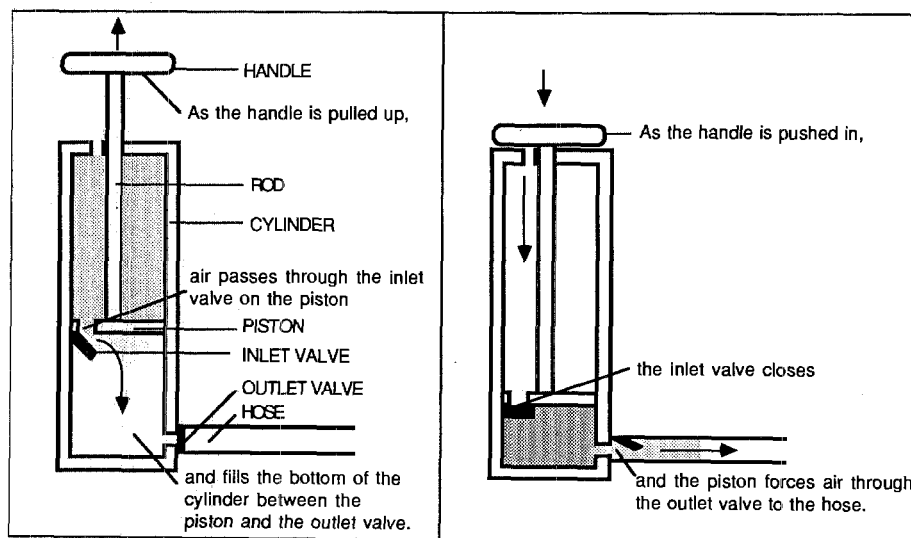


FIGURE 2 Selected illustration and corresponding captions from a multimedia lesson on how a tire pump works. Reprinted from Mayer and Gallini (1990).

the pump example, an acceptable explanation for “why air comes out of the hose” is to say “because you pressed down in the handle.” In describing the phenomenon, no underlying principle or mechanism is involved.

According to the induction-of-rules approach, a scientific explanation exists when a scientist notices relations among events and induces a rule. In the pump example, you might notice that when you push the handle slightly, only a little air comes out of the hose but when you press the handle all the way down a lot of air comes out. Based on these observations you may induce the rule, “the harder the press on the handle, the more air comes out of the hose.” In its most refined form, the induced rule may be expressed mathematically in the form of an equation. Again, no underlying principle or mechanism is involved.

According to the induction-of-models view, an explanation involves the construction of a mechanistic model in which a change in the state of one part is related to a change in the state of another part in a principled way. In the pump example, the explanation for why air goes out of the hose is that the air pressure is greater inside the cylinder than outside. Following the work of Craik (1943), Johnson-Laird (1989, p. 467) argues that explanation occurs when “human beings translate external events into internal models.” The result is a mental model, a “dynamic representation or simulation of the world” (Johnson-Laird, 1989, p. 467).

Which knowledge is needed for understanding an explanation: precise descriptions, formal rules, or mechanistic models? For purposes of our work, we have embraced the invention-of-models view of scientific explanation. In short, we propose that understanding a scientific explanation involves the construction of a mental model of the to-be-explained system. For example, Bromage and Mayer (1981) found that students who were successful in solving problems based on reading a passage on how cameras work tended to remember information about how a change in one part affected a change in another part rather than descriptions or rules. In this case, understanding—as measured by problem-solving transfer—was related to the learner’s mental model of the camera—as measured by recall of a cause-and-effect mechanism.

Analogies and metaphors can be useful in helping learners to construct mental models (Gentner & Stevens, 1983; Halpern, Hansen, & Riefer, 1990; Mayer, 1989a; Vosniadou & Ortony, 1989). In this review, we examine the instructional hypothesis that mental model construction can be enhanced when appropriate models are presented visually as well as verbally. Unlike other studies (e.g., see Schnotz & Kulhavy, 1994) that focus on the role of visual and verbal representations as aids to remembering (which can be measured by tests of retention), this review focuses on the role of visual and verbal representations as aids to understanding (which can be measured by tests of problem-solving transfer).

## A GENERATIVE THEORY OF MULTIMEDIA LEARNING

At this time, the technology for multimedia education is developing at a faster pace than a corresponding science of how people learn in multimedia environments. Technological advances in computer-based graphics—including animation—and text-based graphics—including the use of illustrations—have not been matched by corresponding scientific advances in understanding how people learn from pictures and words (Reiber, 1990).

Our research is motivated by the idea that the design of multimedia instructional materials should be based on a theory of meaningful learning. In particular, our goal is to understand how people integrate verbal and visual information during multimedia learning. In this section I propose a generative theory of multimedia learning that draws on Wittrock’s (1974, 1989) generative theory (as well as extensions by Mayer, 1984, 1993c; and Sternberg, 1985), and Paivio’s (1986; Clark & Paivio, 1991) dual coding theory (as well as extensions by Baddeley, 1992; Mayer, 1992a, 1993a, 1993b; and Schnotz, 1993b). From generative theory, I take the idea that meaningful learning occurs when learners *select* relevant information from what is presented, *organize* the pieces of information into a coherent mental representation, and *integrate* the newly constructed representation with others. From dual coding theory, I take the idea that these cognitive processes occur within two separate information processing systems: a *visual* system for processing visual knowledge and a *verbal* system for processing verbal knowledge.

In a generative theory of multimedia learning, the learner is viewed as a knowledge constructor who actively selects and connects pieces of visual and verbal knowledge. The basic theme of a generative theory of multimedia learning is that the design of multimedia instruction affects the degree to which learners engage in the cognitive processes required for meaningful learning within the visual and verbal information processing systems.

Figure 3 summarizes some cognitive conditions for the construction of meaningful learning in a multimedia environment: selecting words and selecting images from the presented material, organizing words and organizing images into coherent mental representations, and integrating the resulting verbal and visual representations with one another. The first constructive process involves paying attention to relevant aspects of visual and verbal information that has entered the information processing system through the eyes and ears. Mayer (1984, p. 32) refers to this process as “selecting,” which involves “selecting information from the text and adding that information to working memory,” whereas Sternberg (1985, p. 107) refers to this process as “selective encoding,” which involves “sifting out of relevant and irrelevant information.” Aspects of Paivio’s (1986, p. 69) “representational processing” are involved in the selecting process

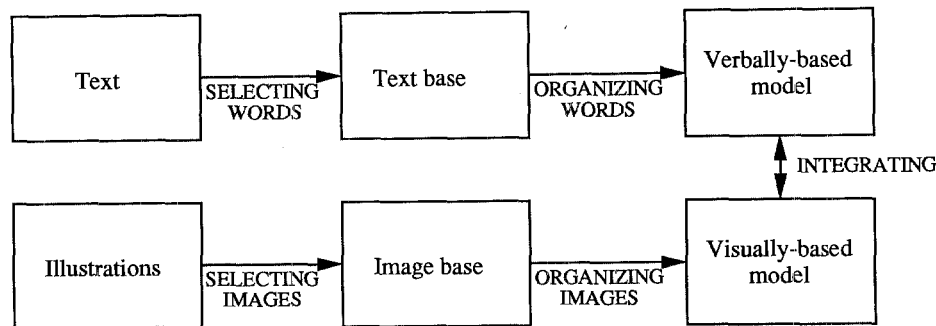


FIGURE 3 A generative model of multimedia learning.

in that verbal stimuli activate the construction of a verbal representation and visual stimuli activate the construction of a visual representation. In models of human information processing, the process of selecting relevant information involves taking information from the sense receptors into visual and verbal short-term memory.

Within the context of a dual-coding view, there can be two kinds of selecting processes: selecting words and selecting images. From the verbal information that is presented, the learner selects relevant words for a verbal representation. This process is indicated by the “selecting words” arrow in Figure 3 and results in the construction of a propositional representation or text base. Similarly, from the visual information that is presented, the learner selects relevant images for a visual representation. This process is indicated by the “selecting images” arrow shown in Figure 3 and results in the construction of what can be called a pictorial representation or image base.

Once the learner has selected visual and verbal material for processing in visual and verbal short-term memory, respectively, the next step is to organize the selected material in a more coherent way. Mayer (1984, p. 32) calls this process “organizing” and describes it as “organizing the selected information in working memory into a coherent whole”; Sternberg (1985, p. 107) calls this process “selective combination” and describes it as “combining selectively encoded information in such a way as to form an integrated ... internally connected whole.” Aspects of Paivio’s (1986, p. 69) “associative processing” are involved in the process of organizing in that associations are formed among elements within the visual information processing system or within the verbal information processing system. In models of human information processing, organizing involves a transformation of verbal knowledge within verbal short-term memory and a transformation of visual knowledge within visual short-term memory.

The “organizing words” arrow in Figure 3 indicates that the learner reorganizes the text base into what can be called a verbal mental model of the situation described in the text. This transformation takes place within verbal short-term memory. Similarly, the “organizing images” arrow indicates that the

learner reorganizes the image base into what can be called a visual mental model of the situation depicted in the pictures. In summary, within the visual and verbal information processing systems respectively, the learner constructs what has been called a situation model—a coherent mental representation of a system in which the parts are related to one another in logical ways (Greeno, 1989).

After constructing a verbally based model and a visually based model, the learner’s final step is to build connections between the two representations. This process is called “integrating” by Mayer (1984, p. 33) and is described as “connecting the organized information to other familiar knowledge structures already in memory.” Similarly, Sternberg (1985, p. 107) uses the term “selective comparison” to refer to the process of “relating newly acquired or retrieved information ... to old knowledge so as to form an externally connected whole.” This process is related to Paivio’s (1986, p. 69) “referential processing” in which connections are between representations in the verbal and visual information processing systems. In models of human information processing, this process, which Paivio (1986, p. 87) calls “building referential connections,” is represented by an arrow between visual short-term memory and verbal short-term memory.

In Figure 3, the “integrating” arrow refers to the process of building one-to-one correspondences between the verbal and visual representations of the material. This process takes place within short-term memory (or a portion of it called working memory). For the integrating process to take place, the visual information must be held in visual short-term memory at the same time that the corresponding verbal information is held in verbal short-term memory. However, the holding capacity of short-term memory is limited, so integrating visual and verbal information during learning is constrained by memory load (Baddeley, 1992; Chandler & Sweller, 1991; Sweller, Chandler, Tierney, & Cooper, 1990).

As an example, consider the cognitive processes involved in learning from a textbook lesson (or an electronic encyclopedia entry) on pumps that contains text and illustrations (or narration and animation) explaining how a bicycle tire pump works. The first step for the constructive learner is to select the relevant words from the presented prose, including the

names of the key parts of the pump (e.g., handle, piston, inlet valve, outlet valve, lower part of the cylinder, hose) and the major actions (e.g., the handle is pulled up, the piston moves, the inlet valve opens, and so on). Correspondingly, the learner must select relevant images from the illustrations or animations, including images of the main parts (e.g., the handle, piston, etc.) and actions (e.g., the handle moving from the down to the up position, the valve moving from the closed to the open position, and so on) of the pump. The left side of Figure 4 provides examples of the verbal information that must be selected; the right side provides examples of the visual information that must be selected.

Then, the selected verbal and visual knowledge must be organized into coherent representations. The process of organizing words involves putting the actions into a cause-and-effect chain, by constructing causal relations among events. For example, "the handle is pulled up" causes "the piston moves up" which causes "the inlet valve opens" and "the outlet valve closes" and so on. Similarly, the process of organizing images involves fitting the images into a cause-and-effect chain, in which the change in one image is the cause of the change in the next one. For example, the image of the handle moving from the down to the up position is the cause of the image of the piston moving from the down to the up position, which is the cause of the inlet valve moving from the closed to the open position and the outlet valve moving from the open to the closed position, and so on.

Finally, the process of integrating occurs when the learner builds one-to-one mappings between the elements, actions, and causal relations in the verbal and visual representations. For example, in mapping elements, the learner must realize that the word *handle* corresponds to the image of the handle at the top of the pump, that the word *piston* refers to the image of the circular object that can slide inside the cylinder, the word *inlet valve* is the same as the image of the one-way door that attached to the piston, and so on. In mapping actions, the learner notes that the statement "the handle is pulled up" corresponds to the image of the handle moving from the down to the up position, the statement "the piston moves up" corresponds to the image of the handle moving from the down to the up position within the cylinder, the statement "the inlet valve opens" refers to the image of the inlet valve moving from the closed to the open position, and so on. In mapping relations, the learner finds that the causal relations between actions stated as words correspond to the causal relations between actions depicted as images; for example, the learner must realize that the cause-and-effect relation between "the handle is pulled up" and "the piston moves up" is the same as the causal relation between the image of the handle moving from the down to the up position and the image of the piston moving from the down to the up position within the cylinder. In summary, meaningful learning occurs when a learner builds coherent mental representations of a cause-and-effect system in verbal short-term memory and in visual short-term

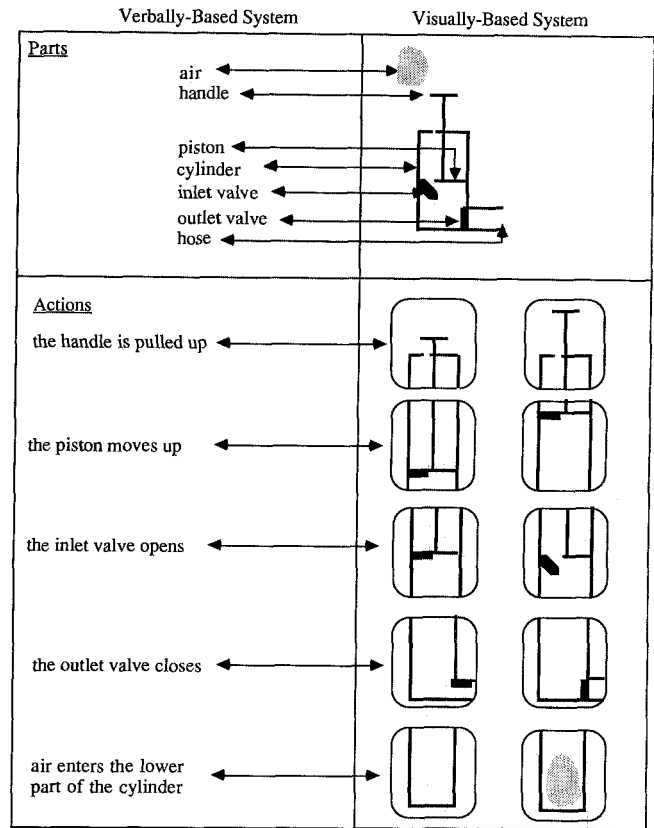


FIGURE 4 Verbal and visual representations of how a tire pump works. Reprinted from Mayer and Sims (1994).

memory, and builds systematic connections between the verbal and visual representations.

I derive several predictions of this generative theory of multimedia learning concerning (a) whether or not multimedia instruction is effective, (b) when multimedia instruction is effective, and (c) for whom multimedia instruction is effective. The predictions are based on the idea that learners are best able to build connections between verbal and visual representations when corresponding components of the verbally based model and visually based model are in short-term memory at the same time. This would occur, for example, when the verbal representation of an action and the visual representation of the same action are simultaneously in the learner's short-term memory. Furthermore, the predictions are based on the idea that this situation is more likely to occur when the text and illustrations are presented contiguously on the page rather than separately or when the narration and animation are presented simultaneously rather than successively. The following sections spell out the predictions and review research findings. First, a classic question that dominated early media research is briefly examined and reframed; then, three current questions based on the predictions of the model are addressed in detail; and finally, a question suggested for future research is briefly reviewed.

## RESEARCH ON MULTIMEDIA LEARNING

### Question 1 (Media Effects): Is One Medium Better Than Another?

A persistent, if somewhat unproductive, question in media research concerns whether one medium is more effective than another (Clark, 1983; Clark & Salomon, 1986; Salomon, 1979/1994; Wetzel et al., 1994). For example, in the domain of multimedia learning, a version of this question is: "Are computers more effective than textbooks?" To answer this question, one could compare the consequences of teaching a lesson using a textbook that contains words and illustrations versus teaching the same lesson using computer-generated graphics and narration.

Although this media question has never been the focus of our research, we have reanalyzed our studies in an attempt to test for media effects. Overall, we have conducted four studies (Mayer, 1989b, Experiments 1 and 2; Mayer & Gallini, 1990, Experiments 1 and 2) in which students read a passage about braking or pumping systems that contained text coordinated with illustrations and five studies (Mayer & Anderson, 1991, Experiments 1 and 2; Mayer & Anderson, 1992, Experiments 1 and 2; Mayer & Sims, 1994, Experiment 2) in which students learned the same material by viewing a computer-generated animation with concurrent narration. In these studies the mean number of creative solutions on a subsequent problem-solving test was 7.5 for students in the computer-based medium and 7.3 for students in the book-based medium.

Consistent with prior research (Clark & Salomon, 1986), our results do not provide strong evidence of media effects. Overall, computer-based learning seems to yield 3% more solutions on a problem-solving test than does book-based learning—a difference so small as to be inconsequential. Yet, even this conclusion is misleading because, like most studies of media effects, there are serious methodological confounds in comparing the two media. For example, in our studies, the book-based materials contained several hundred words including supplemental (and potentially irrelevant) information, whereas the computer-based narration contained only a summary of the essential information, and students generally read through the book-based material one time, whereas they viewed the computer-based material three times. In short, it is not possible to determine whether media effects are attributable to differences between computer-based and book-based delivery systems or to differences in the content and study conditions of the lessons.

An even more damaging critique of research on media effects concerns the observation that the most important factor in producing cognitive outcomes is not the medium that is used but rather the quality of the instructional message (Clark, 1983, 1994). For example, a closer analysis of our results shows that the same kinds of instructional methods have the same kinds of effects within both media. For example, in both computer-based and book-based media, presenting visual

information alone (i.e., illustrations or animations) does not result in good problem-solving performance. In both media, adding verbal information (i.e., paragraphs or narration) has a strong positive effect on problem-solving performance when the verbal information is coordinated with the visual information (i.e., placed next to it on the page or spoken at the same time it is shown in the animation) but not when they are separated (i.e., the text and illustrations are on different pages or the animation is presented before the narration). In short, it is possible to produce effective and ineffective instruction in both computer-based and book-based media; moreover, in both media, ineffective instruction can be changed into effective instruction by applying the same basic instructional principles (Fleming & Levie, 1993).

A more recent criticism of research on media effects is that the research question is based on an outmoded metaphor of learning (Jonassen, 1991; Mayer, 1992a). By asking which medium is the best for delivering an instructional method, a researcher is assuming that knowledge is a commodity (i.e., a message) that can be put inside someone's head (i.e., delivered). These assumptions underlie a *knowledge acquisition* metaphor of learning in which learning is viewed as the process of adding knowledge to the learner's memory and teaching is viewed as the process of dispensing knowledge to learners. The knowledge acquisition metaphor suggests research on questions about media effects, but does not suggest the kind of questions addressed in the next sections. This view has been largely replaced by a *knowledge construction* metaphor of learning in which learning is viewed as a process of active mental construction within the learner and teaching is viewed as a process of fostering and supporting students' efforts to construct their own knowledge. The model of learning presented in Figure 3 is based on a knowledge construction metaphor; it does not suggest questions about media effects but it does suggest questions addressed in the next sections.

In summary, the search for media effects dominated early research on media, but the current consensus among educational psychologists is that questions about the relative effectiveness of various media are no longer productive questions (Clark, 1983, 1994; Clark & Salomon, 1986; Ross, 1994; Salomon, 1979/1994). For example, in a recent debate (Ross, 1994) on Clark's (1983, 1994) assertion that media effects can never be separated from method effects, researchers on all sides agreed that questions regarding media should be reframed. Kozma (1994b, p. 7) argues for research on "media and the methods that employ them, as they interact with the cognitive and social processes by which knowledge is constructed." Jonassen, Campbell, and Davidson (1994, p. 31) call for reframing the debate as "learner-centered rather than media-centered," that is, to focus on how instructional treatments affect cognitive processing within the learner. Clark (1994) calls for research on the cognitive consequences of various instructional methods used within various media rather than on media per se. The rationale for this consensus

is empirical (in general, media effects are small), methodological (in general, it is not possible to separate the effects of media from the effects of the instructional method), theoretical (the driving theoretical issue concerns understanding how people construct meaning from their experiences), and paradigmatic (the knowledge acquisition metaphor has been replaced by the knowledge construction metaphor). For these reasons, this review relegates media-effects questions to the past and focuses on what I hope are more productive questions in the remainder of this review. Similarly, in a recent analysis of research on media effects Kozma (1994a, p. 13) noted that "it is time to shift the focus of our research from media as conveyors of methods to media and methods as facilitators of knowledge-construction and meaning-making on the part of learners." This is the approach taken in the following sections.

### Question 2 (Multimedia Effects): Is Multimedia Instruction Effective?

Based on generative theory I predict a multimedia effect in which students who receive a verbal explanation coordinated with a visual explanation (multiple representation group) perform better on problem-solving transfer than students who receive only a verbal explanation (single representation group). In particular, Prediction 1 is that students who receive an explanation in the form of text coordinated with illustrations will generate more creative solutions to problem-solving transfer questions than students who receive an explanation in text form only. Prediction 2 is that students who receive an explanation in the form of narration coordinated with animation will generate more creative solutions to problem-solving transfer questions than students who receive an explanation in narration only.

**Overall results: Multimedia effect.** In a series of eight experiments, my colleagues and I at Santa Barbara compared the problem-solving transfer performance of students who learned from visual and verbal explanations that were coordinated

(multiple representation group) with the performance of students who received only verbal explanations (single representation group). All the studies involved short cause-and-effect explanations of how various devices work, such as brakes, pumps, and generators. Whenever data of individual differences was available, we focused only on students who were classified as low in prior knowledge and high in spatial ability (as explained under Question 4). Three studies compared computer-generated narration coordinated with animation versus computer-generated narration alone, and five studies compared text coordinated with illustrations versus text alone. In all studies students were tested by asking them to generate as many solutions as possible to a series of transfer questions, and a problem-solving transfer score was computed for each student by tallying the number of correct creative solutions generated on the transfer problems. We computed a multimedia effect for each study by dividing the mean transfer score of the multiple representation group by the transfer score of the single representation group and subtracting one from the result.

Table 2 lists the source, topic materials, and size of the multimedia effect for each of three comparisons involving narration and animation and five comparisons involving text and illustrations; Figure 5 presents the same eight comparisons visually as dots along a number line. As can be seen, multimedia effects range from 52% to 110% with a median of 78%. Overall, students who received visual explanations coordinated with verbal explanations produced more than 75% more creative solutions to transfer problems than did students who received the explanation presented only in verbal form. Thus, in each of the eight comparisons there is strong evidence that adding a visual explanation to a verbal one can greatly enhance student understanding, as measured by tests of problem-solving transfer.

**A closer look at Prediction 1: Multimedia effect for narration and animation.** The first prediction was that students exhibit better problem-solving transfer performance

TABLE 2  
Summary of Research on Multimedia Effects

<i>Experiment</i>	<i>Topic</i>	<i>Contiguity Effect</i>
Dual media are narration and animation		
Mayer & Anderson (1991, Experiment 2b)	pump	+110%
Mayer & Anderson (1992, Experiment 1)	pump	+96%
Mayer & Anderson (1992, Experiment 2)	brakes	+97%
Median for narration and animation		+97%
Dual media are text and illustrations		
Mayer (1989b, Experiment 1)	brakes	+65%
Mayer (1989b, Experiment 2)	brakes	+64%
Mayer & Gallini (1990, Experiment 1, low knowledge)	brakes	+87%
Mayer & Gallini (1990, Experiment 2, low knowledge)	pump	+69%
Mayer & Gallini (1990, Experiment 3, low knowledge)	generators	+52%
Median for text and illustrations		+65%



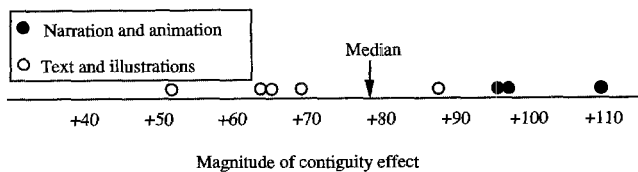


FIGURE 5 Summary of eight tests of the multimedia effect.

when narration is coordinated with animation than when only narration is presented. We tested this prediction in three studies. In our first two studies (Mayer & Anderson, 1991, Experiment 2b; Mayer & Anderson, 1992, Experiment 1) some college students viewed a 30-sec animation depicting changes in a line drawing of a tire pump as the handle was pulled up and pushed down and listened to a concurrent narration using digitized speech that described the actions taking place on the screen (multiple representation group), whereas other college students listened to the same narration without seeing the animation (single representation group). A summary of the materials is presented in Figure 1. For both groups, the materials were presented three times, and then all students were asked to generate as many answers as possible to transfer questions such as shown in Table 1. As predicted, the first and second panels of Figure 6 show that the multiple representation group generated more creative solutions on the transfer problems than the single representation group in each of the experiments, respectively.

In a third study (Mayer & Anderson, 1992, Experiment 2), some college students viewed a 30-sec animation depicting changes in a car's hydraulic braking system when the brake is pressed down and when the brake pedal is let up and listened to a corresponding narration consisting of digitized speech describing the actions in words (multiple representation group), whereas other students listened to the narration without viewing the

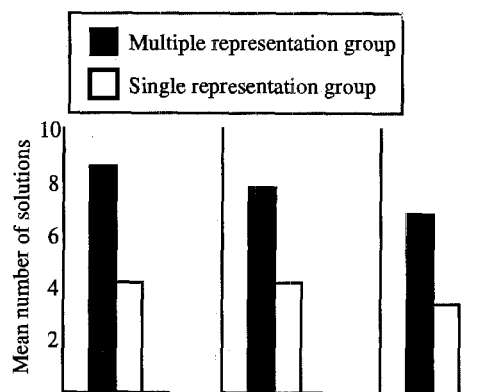


FIGURE 6 Mean number of creative solutions on a problem-solving test for multiple and single representation groups in three experiments involving narration and animation. Frame 1 is adapted from Mayer and Anderson (1991, Experiment 2b), Frame 2 is adapted from Mayer and Anderson (1992, Experiment 1), and Frame 3 is adapted from Mayer and Anderson (1992, Experiment 2).

animation (single representation group). The materials, which are summarized in Figure 7, were presented three times. Then students were asked to generate as many solutions as possible to a series of four transfer problems, such as: "Suppose you press on the brake pedal in your car but the brakes don't work. What could have gone wrong?" The third panel of Figure 6 shows that the multiple representation group generated more solutions than the single representation group.

In three different comparisons, students who received explanations in words coordinated with pictures generated a median of 97% more creative solutions to problems than did students who received explanations only in words.

*A closer look at Prediction 2: Multimedia effect for text and illustrations.* The second prediction was that students exhibit better problem-solving transfer performance when text is coordinated with illustrations than when only text is presented. We tested this prediction in five studies. In our first three studies (Mayer, 1989b, Experiment 1, Experiment 2; Mayer & Gallini, 1990, Experiment 1) some college students read a 750-word passage on braking systems that was accompanied by four captioned illustrations (multiple representation group), whereas other students read the text without illustrations (single representation group). A summary of the materials is presented in Figure 8. Then all students were asked to generate as many answers as possible to transfer questions such as described previously. As predicted, the first, second, and third panels of Figure 9 show that the multiple representation group generated a greater number of creative solutions on the transfer problems than the single representation group in each of the experiments, respectively.

In a fourth study (Mayer & Gallini, 1990, Experiment 2) some college students read a 750-word passage on pumps that contained three captioned illustrations (multiple representation group), whereas other college students read the passage without illustrations (single representation group). The materials, which are summarized in Figure 2, were followed for a 4-item problem-solving test as summarized in Table 1. The fourth panel of Figure 9 shows that the multiple representation group generated more solutions than the single representation group.

In the fifth study (Mayer & Gallini, 1990, Experiment 3), some college students read a 2,000-word passage on electrical generators that contained four captioned illustrations (multiple representation group), whereas other college students read the passage without illustrations (single representation group). Then, all students generated as many solutions as possible to a series of transfer problems, such as: "What can be done to increase the energy output from an electrical generator?" The fifth panel in Figure 9 shows that the multiple representation group generated more solutions than the single representation group.

In five distinct comparisons, students who received explanations in words and corresponding illustrations generated a

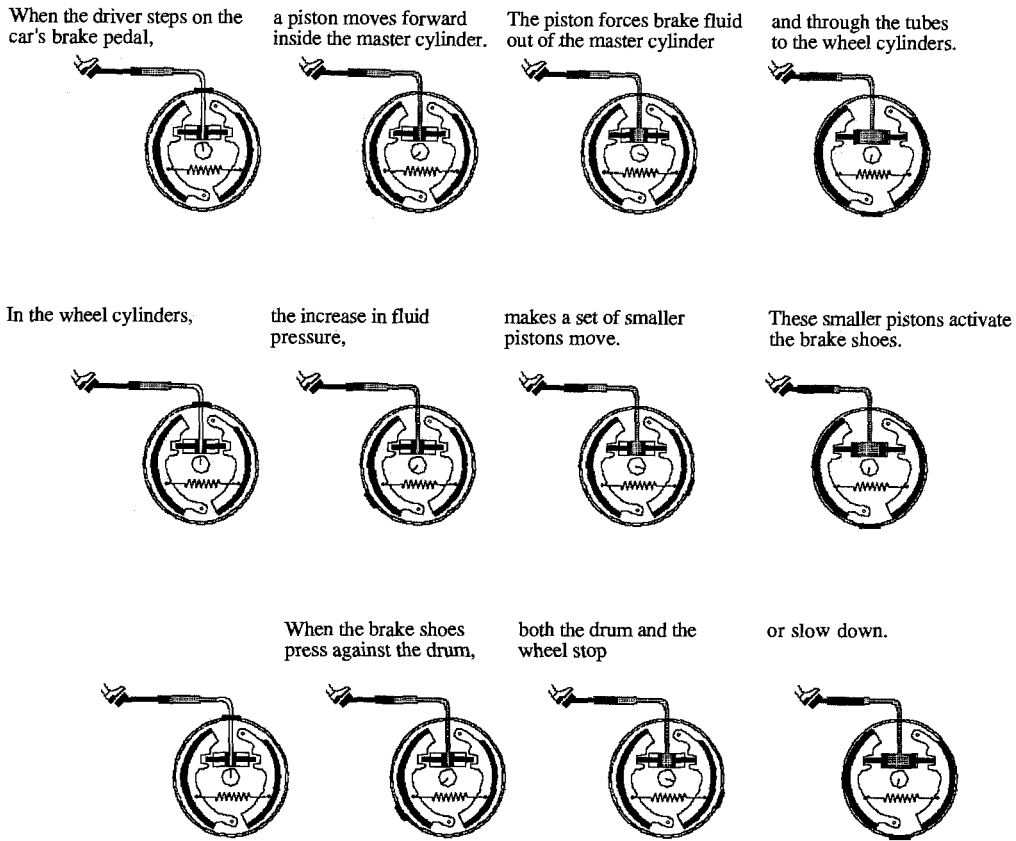


FIGURE 7 Selected animation frames and corresponding narration from a multimedia lesson on how a braking system works. Reprinted from Mayer and Anderson (1992).

median of 65% more creative solutions to problems than did students who received an explanation only in words.

**Question 3 (Contiguity Effects): When Is Multimedia Instruction Effective?**

The foregoing review provides consistent evidence for a multimedia effect: Students' understanding of verbally based explanations can be improved when visually based explana-

tions are added. A reasonable next question concerns the conditions under which multimedia instruction is effective. Based on the generative theory summarized in Figure 3, I predict a contiguity effect in which the problem-solving transfer performance of students who receive words and pictures coordinated with each other (coordinated group) exceeds that of students who receive the same words and pictures separated from one other (separated group). I call this a contiguity effect because multimedia instruction is most effective when words

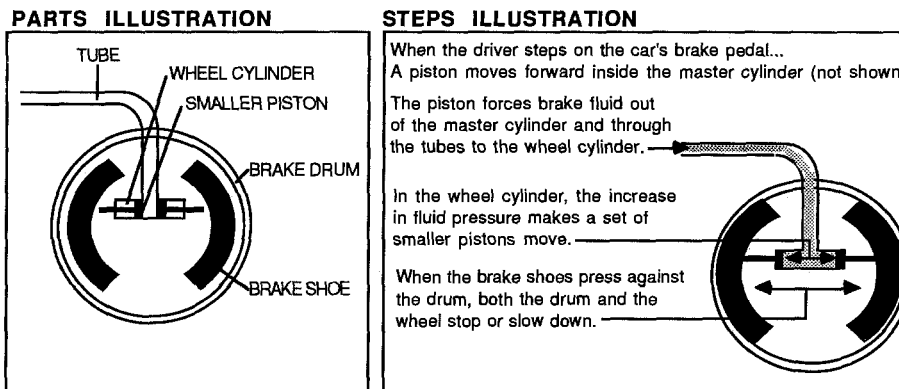


FIGURE 8 Selected illustrations and corresponding captions from a multimedia lesson on how braking systems work. Reprinted from Mayer and Gallini (1990).

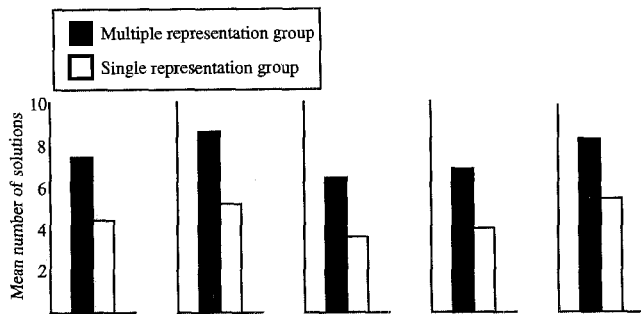


FIGURE 9 Mean number of creative solutions on a problem-solving test for multiple and single representation groups in five experiments involving text and illustrations. Frame 1 is adapted from Mayer (1989b, Experiment 1), Frame 2 is adapted from Mayer (1989b, Experiment 2), Frame 3 is adapted from Mayer and Gallini (1990, Experiment 1), Frame 4 is adapted from Mayer and Gallini (1990, Experiment 2), and Frame 5 is adapted from Mayer and Gallini (1990, Experiment 3).

and pictures are presented contiguously—next to each other on a page for text-based instruction or synchronized in time for computer-based instruction. In particular, this analysis leads to two straightforward predictions: Prediction 3 is that students generate more correct creative solutions to problem-solving transfer questions when text and illustrations are coordinated than when they are separated. Prediction 4 is that students generate more correct creative solutions to problem-solving transfer questions when narration and animation are coordinated than when they are separated.

**Overall results: Contiguity effect.** In a series of 10 experiments, my colleagues and I at Santa Barbara compared the outcomes of multimedia learning when visual and verbal materials were presented simultaneously (coordinated presentation) or separately in time or space (separated presentation). All the studies involved short cause-and-effect explanations of how various systems work, such as braking systems, pumps, electrical generators, the process of light-

ning, and the human respiratory system. Wherever data on individual differences were available, we focused only on students who lacked prior knowledge and who possessed high spatial ability. Six studies used several minutes of computer-generated auditory narration and visual animation: Coordinated presentation involved presenting the narration at the same time as the animation so that changes in the picture on the screen were coordinated with verbal descriptions in the auditory narration, whereas separated presentation involved presenting the narration and animation in succession so that a learner first listens to the narration and then views the animation or vice versa. Four studies used several pages of text and illustrations: Coordinated presentation involved presenting illustrations with verbal captions and labels next to corresponding portions of the text, whereas separated presentation involved presenting the same words and illustrations on separate pages. In all studies students were tested to see the degree to which they could apply the presented explanations to solve new problems, such as writing answers to questions about how to improve the system or how to troubleshoot the system. A problem-solving transfer score was computed for each student by tallying the number of correct creative solutions generated on these transfer problems.

According to the generative theory of multimedia learning, meaningful learning occurs when students select and organize relevant visual and verbal information and systematically integrate the newly constructed visual and verbal representations. Given the resource limitations on working memory, the processes of selecting, organizing, and integrating are more likely to occur when visual and verbal information is presented contiguously rather than separately. To examine this prediction we computed a contiguity effect for each study by dividing the problem-solving transfer score of the coordinated group by the problem-solving transfer score of the separated group and subtracting one from the result.

Table 3 lists the source, topic materials, and size of the contiguity effect for each of six comparisons involving nar-

TABLE 3  
Summary of Research on Contiguity Effects

Experiment	Topic	Contiguity Effect
Dual media are narration and animation		
Mayer & Anderson (1991, Experiment 1)	pump	+45%
Mayer & Anderson (1991, Experiment 2a)	pump	+48%
Mayer & Anderson (1992, Experiment 1)	pump	+70%
Mayer & Anderson (1992, Experiment 2)	brakes	+64%
Mayer & Sims (1994, Experiment 1, high spatial)	pump	+43%
Mayer & Sims (1994, Experiment 2, high spatial)	lungs	+57%
Median for narration and animation		+53%
Dual media are text and illustrations		
Mayer (1989b, Experiment 2)	brakes	+65%
Mayer et al. (1995, Experiment 1)	lightning	+78%
Mayer et al. (1995, Experiment 2, low knowledge)	lightning	+81%
Mayer et al. (1995, Experiment 3)	lightning	+70%
Median for text and illustrations		+74%

ration and animation and four comparisons involving text and illustrations; Figure 10 presents the same 10 comparisons visually as dots along a number line. As can be seen, contiguity effects range from 43% to 81% with a median of 65%. Overall, students who received visual explanations coordinated with verbal explanations produced more than 50% more creative solutions to transfer problems than did students who received the same visual and verbal explanations separated in time or space. Thus, in each comparison there was strong evidence for a contiguity effect consistent with the predictions of a generative theory of multimedia learning.

*A closer look at Prediction 3: Contiguity effect for narration and animation.* Our next prediction was that students would exhibit better problem-solving transfer performance when narration and animation are coordinated than when they are separated. We tested this prediction in six separate studies. In one set of studies (Mayer & Anderson, 1991, Experiments 1 and 2), college students viewed a 30-sec animation depicting changes in a line drawing of a tire pump as the handle is pulled up and pushed down and listened to a narration using digitized speech that described the actions taking place when the handle is pulled up and pushed down. The coordinated group received the animation and narration simultaneously, whereas the separated group received the narration followed by the animation. For both groups, the animation and narration were presented three times, so both groups were exposed to exactly the same materials. Then, all students solved four transfer problems, such as: "Suppose you push down and pull up the handle of the pump several times and no air comes out. What could have gone wrong?" As predicted, the first and second panels of Figure 11 show that the coordinated group generated more creative solutions on the transfer problems than the separated group in each of the experiments, respectively.

Although these findings are consistent with our predictions, the studies can be criticized on methodological grounds. First, the separated group received three presentations of the narration before animation (NANANA in which N represents narration and A represents animation). Would the same results be obtained if students received three presentations of animation before narration (ANANAN), three presentations of narration followed by three presentations of animation (NNNAAA), or three presentations of animation followed by three presentations

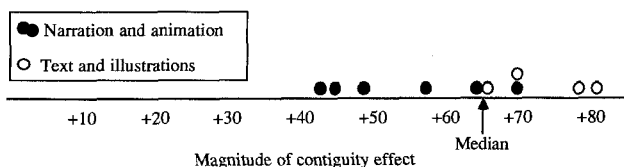


FIGURE 10 Summary of 10 tests of the contiguity effect in multimedia learning.

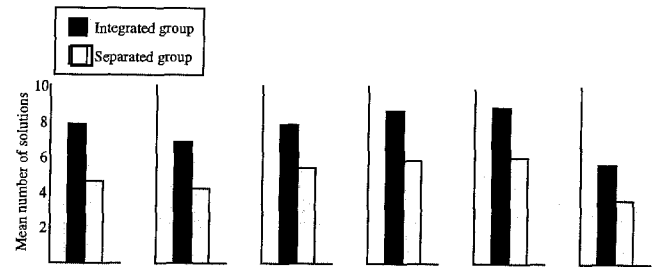


FIGURE 11 Mean number of creative solutions on a problem-solving test for integrated and separated groups in six experiments involving narration and animation. Frame 1 is adapted from Mayer & Anderson (1991, Experiment 1), Frame 2 is adapted from Mayer and Anderson (1991, Experiment 2a), Frame 3 is adapted from Mayer and Anderson (1992, Experiment 1), Frame 4 is adapted from Mayer and Anderson (1992, Experiment 2), Frame 5 is adapted from Mayer and Sims (1994, Experiment 1), and Frame 6 is adapted from Mayer and Sims (1994, Experiment 2).

of narration (AAANNN)? Second, in our original study we had focused only on pumps. Would the same pattern of results hold if we used new materials such as braking systems? We conducted two additional studies (Mayer & Anderson, 1992, Experiments 1 and 2) to examine these methodological issues. The first study was like the previous ones except that four versions of separated group were used: ANANAN, NANANA, NNNAAA, and AAANNN. In the second study, college students viewed a 30-sec animation depicting changes in a car's hydraulic braking system when the brake pedal is pressed down and when the brake pedal is let up and listened to a narration consisting of digitized speech describing the actions taking place inside the system when the pedal is pressed down and let up. Selected frames of animation and the corresponding narration on brakes are given in Figure 7. Students in the coordinated group received three presentations of the narration and animation simultaneously, whereas students in the separated group received one of the four versions described previously. Then, students were asked to solve four transfer problems; for example, students who learned about brakes were asked: "Suppose you press on the brake pedal in your car but the brakes don't work. What could have gone wrong?" In both studies, there were no significant differences among the four versions of the separated group, so they were combined for presentation in the third and fourth frames of Figure 11, respectively.

In the next phase of our research we compared coordinated and separated presentation of narrations and animations of pumping systems and the human respiratory system (Mayer & Sims, 1994, Experiments 1 and 2). The first study on pumps was similar to previous studies; the second study on how the lungs work used the same procedure but involved a 45-sec animation and 100-word narration describing what happens when air is inhaled and exhaled. Selected frames from the animation and the corresponding narration are presented in Figure 12. As in previous studies, after instruction students



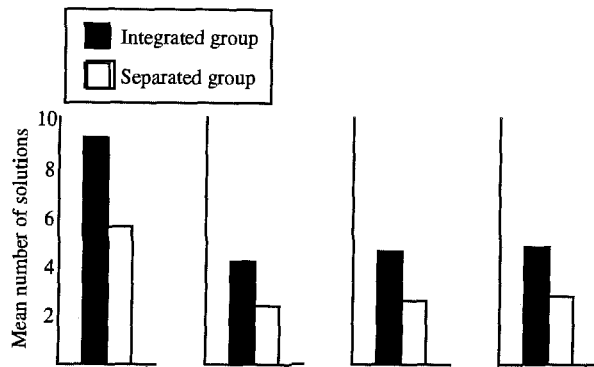


FIGURE 13 Mean number of creative solutions on a problem-solving test for integrated and separated groups in four experiments involving illustrations and text. Frame 1 is adapted from Mayer (1989b, Experiment 2), and Frames 2, 3, and 4 are adapted from Mayer, Steinhoff, Bower, and Mars (1995, Experiment 1, 2, and 3, respectively).

Further tests of the first prediction can be found in a recent study on learning about lightning from a textbook lesson (Mayer, Steinhoff, Bower, & Mars, 1995). Students in the coordinated presentation group were given a passage on lightning that contained 600 words and five annotated illustrations placed near their corresponding paragraphs. The words explained the cause-and-effect chain of events that leads to the creation of lightning; the annotated illustrations depicted each of five major events in the formation of lightning and included short captions and labels taken verbatim from the text. Figure 14 shows the five annotated illustrations. In contrast, students in the separated presentation group were given the 600-word text on one page (for the same amount of time that the coordinated group viewed their coordinated materials), and then, after the text page was taken away, they were given another page containing the five illustrations without any words. Thus, both groups read the same words and viewed the same graphics, but the words and pictures were coordinated for one group and separated for the other. The transfer test asked students to generate solutions to four problems such as: "What could you do to decrease the intensity of a lightning storm?" or "Suppose you see clouds in the sky, but no lightning. Why not?" The second, third, and fourth panels of Figure 13 summarize the mean number of correct creative solutions generated by students in the coordinated and separated groups across the three studies. In each case, students who received words and illustrations in coordinated fashion generated substantially more creative problem solutions than did students who received the same words and illustrations in separated fashion for an even longer period of study time.

In four studies, students who received coordinated presentation of text and illustrations produced a median of 74% more correct creative solutions to problems than did students who received exactly the same information in separated form.

#### Question 4 (Interaction Effects): For Whom is Multimedia Instruction Effective?

The foregoing two sections provide consistent evidence for multimedia and contiguity effects, respectively. A reasonable next question concerns the role of individual differences in the effectiveness of multimedia instruction. In this section, I derive and test several predictions concerning individual differences in prior knowledge and spatial ability based on generative theory. Prediction 5 is that multimedia effects and contiguity effects are strong for low prior knowledge learners but not for high prior knowledge learners. The rationale for this prediction is that high prior knowledge learners are more able than low prior knowledge learners to generate mental images as they read text or listen to narration so they are more likely to have verbal and visual representations in short-term memory at the same time. In this case prior knowledge in a domain, such as knowing how pistons and valves work, can compensate for lack of coordination of words and pictures during instruction.

Prediction 6 is that these multimedia and contiguity effects are strong for high spatial ability learners but not for low spatial ability learners. The rationale for this prediction is that low spatial ability learners must devote more cognitive resources to constructing mental images than do high spatial ability learners, so even when pictures and words are presented in a coordinated fashion low spatial ability learners are less likely to have cognitive resources available to build connections between words and pictures than are high spatial ability learners. In this case, spatial ability can enhance the coordination of words and pictures during instruction.

*Overall results: Attribute  $\times$  Treatment interactions (ATIs).* An ATI occurs when the effect of a treatment (such as coordinated presentation of visual and verbal explanations) depends on the characteristics of the learner. Of the studies of multimedia effects summarized in Table 2, three involve ATIs based on prior knowledge (Mayer & Gallini, 1990, Experiments 1, 2, and 3). In each study, multimedia effects are strong for students who lack prior knowledge about mechanical devices—87%, 69%, and 52%, respectively—whereas the effects are weak for high prior knowledge students—15%, 8%, and 11%, respectively. Of the studies of contiguity effects summarized in Table 3, one involves an ATI based on prior knowledge in which the contiguity effect is strong for low prior knowledge learners—81%—but weak for high prior knowledge learners—0% (Mayer et al., 1995, Experiment 2). Of the studies of contiguity effects summarized in Table 3, two involve an ATI based on spatial ability (Mayer & Sims, 1994, Experiment 1 and Experiment 2). In each case, the contiguity effect is strong for high spatial students—43% and 57%, respectively—but not for low spatial students—7% and 0%, respectively. These results are consistent with Predictions 5 and 6.

### The Process of Lightning

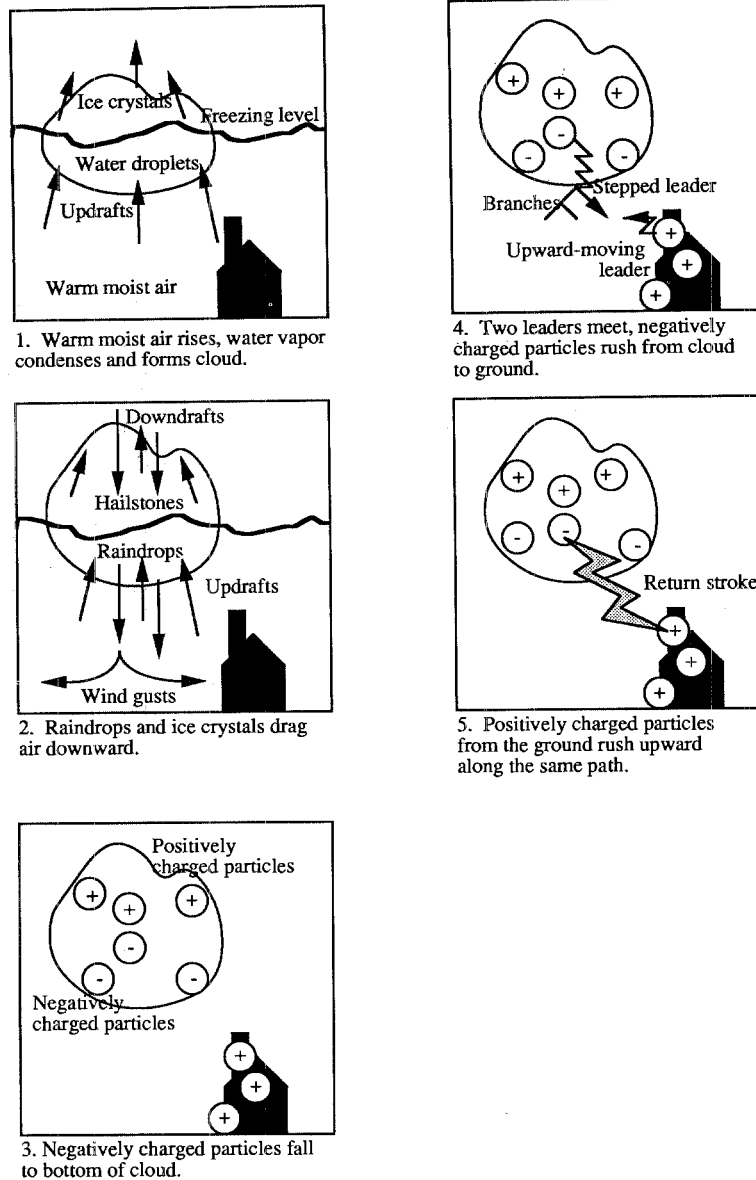


FIGURE 14 Selected illustrations and corresponding captions from a multimedia lesson on how lightning storms develop. Reprinted from Mayer et al. (1995).

*A closer look at Prediction 5: Multimedia and continuity effects for low prior knowledge learners but not for high prior knowledge learners.* A generative theory of multimedia learning also makes predictions concerning the role of individual differences in learning from visual and verbal materials. First, we predict that the multimedia and contiguity effects will be stronger for low prior knowledge learners than for high prior knowledge learners. Students who possess high levels of prior knowledge will be more likely than low prior knowledge learners to create their own mental images as the verbal explanation is presented and

thus to build connections between verbal and visual representations. In contrast, students who lack prior knowledge will be less likely than high prior knowledge learners to independently create useful mental images solely from the verbal materials. Thus, low prior knowledge learners are more likely than high prior knowledge learners to benefit from the contiguous presentation of verbal and visual explanations.

This prediction concerning contiguity effects was examined in one study involving text and illustrations on how lightning works (Mayer et al., 1995, Experiment 2). In order

to assess experience, college students were asked to rate their knowledge of meteorology on a 5-point scale (from *very little* to *very much*) and to place a check mark next to each of the statements that applied to them (such as: "I regularly read the weather maps in a newspaper" or "I know what a cold front is"). Students who rated their knowledge as *very little* and who checked fewer than three items were classified as low prior knowledge learners, whereas students who rated their knowledge as above *very much* and who checked three or more items were classified as high prior knowledge. Consistent with the prediction, a significant contiguity effect was obtained for low prior knowledge learners but not for high prior knowledge learners. The first panel of Figure 15 shows that low prior knowledge students who received integrated instruction generated more creative solutions on the problem-solving test than low prior knowledge students who received separated instructions, whereas for high prior knowledge students there was no significant difference between the integrated and separated groups.

Similarly, in three additional studies of the multimedia effect involving text and illustrations on pumps, brakes, and electrical generators, low prior knowledge learners who received coordinated text and illustrations performed much better on tests of problem-solving transfer than did those who received text without illustrations, whereas high prior knowledge learners performed well under both treatments (Mayer & Gallini, 1990, Experiments 1, 2, and 3). These results are summarized in the second, third, and fourth panels of Figure 15, respectively.

A closer look at Prediction 6: *Multimedia and continuity effects for high spatial ability learners but not for low spatial ability learners.* A second dimension of individual differences in multimedia learning involves spatial ability. In particular, I predict that the contiguity effect will be stronger for high spatial ability learners than for low spatial ability learners. Students who possess low levels of spatial ability may be less able than high spatial ability learners to take advantage of contiguous presentation of visual and verbal

material because they have more difficulty in holding and manipulating the visual representation in memory, as is required to integrate the verbal and visual representations. In contrast, students who possess high levels of spatial ability may be more likely than low spatial ability learners to benefit from contiguous presentation because they are more facile at holding and manipulating visual representations in memory, as is required for integrating the verbal and visual representations. This prediction was examined in two studies involving narration and animation (Mayer & Sims, 1994, Experiments 1 and 2), in which spatial ability was measured by paper-and-pencil tests of mental rotation and paper folding. Consistent with the prediction, in each case, a significant contiguity effect was obtained for high spatial ability learners but not for low spatial ability learners. These ATIs are summarized in the two panels of Figure 16, respectively.

The ATIs in this section concerning prior knowledge and spatial ability temper the multimedia and contiguity effects reported in the previous sections. In particular, the multimedia and contiguity effects reported in the previous sections seem to occur mainly for low-knowledge students and high spatial ability students. Thus, well-designed multimedia representations are most helpful for learners who lack prior knowledge and who possess high spatial ability.

#### Question 5 (Split-Attention Effects): In Multimedia Learning, Is Narration More Effective Than Text?

So far I have examined three current questions: Is multimedia effective?, When is multimedia effective?, and For whom is multimedia effective? and one that belongs to the past: Is one medium more effective than another? In this section, I examine a question that belongs to the future, concerning the format of verbally presented information.

In a pilot study conducted by Matt Mendrala, students viewed an animation depicting a car's braking system and answered problem-solving questions, similar to studies de-

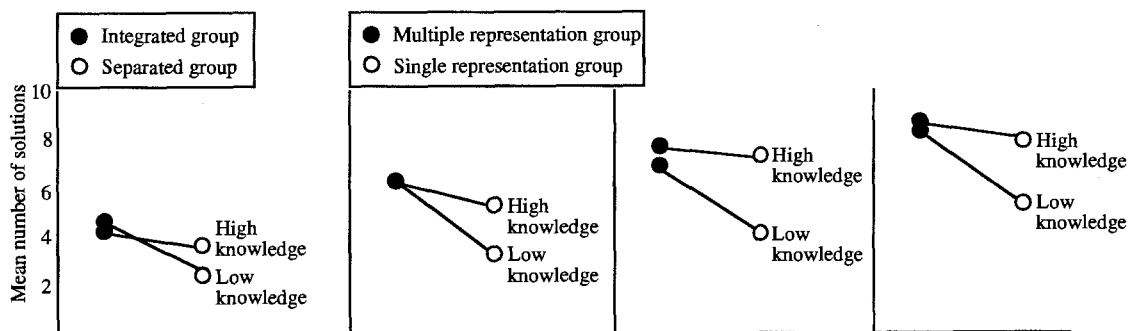


FIGURE 15 Mean number of creative solutions on a problem-solving test for integrated and separated groups (and for multiple and single representation groups) by level of prior knowledge in four experiments. Frame 1 is adapted from Mayer et al. (1995, Experiment 2), and Frames 2, 3, and 4 are adapted from Mayer and Gallini (1991, Experiment 1, 2, and 3, respectively).



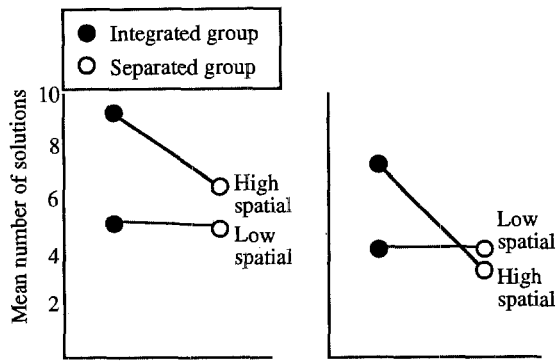


FIGURE 16 Mean number of creative solutions on a problem-solving test for integrated and separated groups by level of spatial ability in two experiments. Frames 1 and 2 are adapted from Mayer and Sims (1994, Experiment 1 and 2, respectively).

scribed in previous sections. Some students heard a concurrent narration describing the cause-and-effect process in words (animation-with-narration group) whereas others saw the same words as text on the screen presented concurrently with the animation (animation-with-text group). According to the model shown in Figure 3, students in both of these groups should be likely to perform each of the three cognitive processes required for meaningful learning: construct a verbal representation, construct a visual representation, and construct connections between the two representations. However, according to Chandler and Sweller's (1991; Sweller et al., 1990) split-attention theory and Baddeley's (1992) model of working memory, verbal information presented visually may be processed differently than verbal information presented acoustically. In particular, when text and animation are both presented visually, the learner's visual attention must be split between the animation and the text. When visual attention is overloaded, some of the information may be lost and the process of constructing connections between visual and verbal information will be disrupted. In contrast, when text is presented auditorily and the corresponding animation is presented visually, the learner can process the representation of the text within an acoustic working memory and the representation of the animation within a visual working memory, which reduces the load on visual attention. This situation increases the chances that the learner will be able to construct connections between verbal and visual representations of the causal chain.

As predicted, students in the animation-with-narration group produced approximately 50% more creative solutions on problem-solving questions than did students in the animation-with-text group. These results are consistent with similar results by Chandler and Sweller (1991) and suggest an important refinement in the processing model in Figure 3. In particular, the way that verbal information is presented can influence the demands placed on acoustic and visual processing resources within the human information processing system. These preliminary results point to the need for future

research concerning the role of cognitive load within the learner's visual and acoustic information processing systems. These preliminary results also suggest future practical implications concerning the advantages of narration over text in the design of multimedia programs. Similarly, Paas and Van Merriëboer (1994, p. 367) noted that "one of the main problems in conventional instructional design for complex tasks is the lack of control of cognitive load" and suggested techniques for reducing cognitive load and guiding learners' attention. For example, students who read summaries of scientific processes perform better on problem-solving transfer than do students who read full lessons, presumably because summaries reduce cognitive load and direct the learner's attention (Mayer, Bove, Bryman, Mars, & Ta-pangco, 1996).

## CONCLUSION

The potential for computer-based aids to learning remains high, although the current contribution of technology to pedagogic innovation is frustratingly low. Instructional development is too often based on what computers can do rather than on a research-based theory of how students learn with technology. In particular, the visual-based power of computer technology represents a grossly underutilized source of potential educational innovation. In computer-based multimedia learning environments students have the opportunity to work easily with both visual and verbal representations of complex systems, but to fruitfully develop these potential educational opportunities, research is needed in how people learn with multimedia.

The program of research on multimedia learning described in this report involves a search for useful research questions. The traditional question concerning the effectiveness of one medium over another proved to be an unproductive question, as have previous studies on media effects. An analysis of a cognitive theory of verbal and visual knowledge construction yielded several productive questions concerning the role of visual and verbal modes of presenting scientific explanations: questions about multimedia effects, contiguity effects, interaction effects, and split-attention effects. In short, this review demonstrates that progress in multimedia research depends partly on searching for the right kinds of questions.

Overall, we can derive some useful theoretical and practical implications from our research on multimedia learning. On the theoretical side, the results of an extensive series of studies using a variety of materials provide consistent support for the generative theory of multimedia learning. According to this theory, coordinated presentation of explanative words and pictures is effective because it helps guide learners' cognitive processes. First, the captioned illustrations and narrated animations serve as signals that help learners select relevant visual and verbal information. Second, captioned illustrations and narrated animations serve as organizers that help learners build cause-and-effect relations among the

pieces of verbal information and among the pieces of visual information. Third, the captioned illustrations and narrated animations serve as coordinators that help learners build one-to-one connections between actions in the visual representation and in the verbal representation. Each of these proposed cognitive functions of captioned illustrations and narrated animations should be considered as a research hypothesis requiring further study.

On the practical side, our work pinpoints an instructional manipulation that increases the generation of creative problem solutions by an average of more than 50%. For textbook design, our results point to the importance of including what I have called explanative illustrations (Mayer, 1993b): a series of two or more frames that show the state of each crucial part of the system at various points, such as the pump when the handle is pushed down and when the handle is pulled up. However, presenting the illustrations alone is insufficient to improve problem-solving performance. In addition, the illustrations need captions that summarize the major actions depicted in the illustration and verbal labels pointing to the major parts described in the caption. For software design, our results point to the importance of coordinating an auditory narration with a visual animation of the step-by-step changes in the system. Presenting an animation—however clever—without concurrent narration is unlikely to promote meaningful learning.

In our studies several conditions may have contributed to our results: the materials, the learners, and the tests. First, we used visual and verbal material that explained how something worked. In short, the material was potentially meaningful because it could be understood as a cause-and-effect system. Had we focused on descriptive passages that presented lists of facts, it is doubtful that we could have obtained a contiguity effect. Second, we focused mainly on students who lacked prior knowledge and were not poor in spatial ability. Had we tested only high prior learners, it is doubtful that we could have obtained a contiguity effect. Third, we used tests intended to measure meaningful learning, namely tests of problem-solving transfer. Had we focused solely on retention of the presented material, we may not have obtained a contiguity effect. In summary, contiguous presentation of visual and verbal material may be most important when the material is a cause-and-effect explanation of a simple system, when the learners are inexperienced, and when the goal is meaningful learning.

#### ACKNOWLEDGMENTS

This article is based on the Division 15 Presidential Address at the 1992 annual convention of the American Psychological Association.

#### REFERENCES

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

- Bromage, B. K., & Mayer, R. E. (1981). Relationship between what is remembered and creative problem-solving performance in science learning. *Journal of Educational Psychology*, 73, 451–461.
- 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.
- Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, 53, 445–459.
- Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, 42(2), 21–30.
- Clark, R. E., & Salomon, G. (1986). Media in teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 464–478). New York: Macmillan.
- Craik, K. (1943). *The nature of explanation*. Cambridge, England: Cambridge University Press.
- Fleming, M., & Levie, W. H. (Eds.). (1993). *Instructional message design: Principles from the behavioral and cognitive sciences* (2nd ed). Englewood Cliffs, NJ: Educational Technology Publications.
- Gentner, D., & Stevens, A. L. (Eds.). (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Greeno, J. G. (1989). Situation models, mental models, and generative knowledge. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 285–318). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Halpern, D. F., Hansen, C., & Riefer, D. (1990). Analogies as an aid to understanding and memory. *Journal of Educational Psychology*, 82, 298–305.
- Houghton, H. A., & Willows, D. M. (Eds.). (1987). *The psychology of illustration: Vol. 2, Instructional issues*. New York: Springer-Verlag.
- Johnson-Laird, P. N. (1989). Mental models. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 469–499). Cambridge, MA: MIT Press.
- Jonassen, D. H. (1991). Objectivism versus constructivism: Do we need a new philosophical analysis? *Educational Technology Research and Development*, 39(3), 5–14.
- Jonassen, D. H., Campbell, J. P., & Davidson, M. E. (1994). Learning with media: Restructuring the debate. *Educational Technology Research and Development*, 42(2), 31–39.
- Kozma, R. B. (1991). Learning with media. *Review of Educational Research*, 61, 179–212.
- Kozma, R. B. (1994a). A reply: Media and methods. *Educational Technology Research and Development*, 42(3), 11–14.
- Kozma, R. B. (1994b). Will media influence learning? Reframing the debate. *Educational Technology Research and Development*, 42(2), 1–19.
- Mandl, H., & Levin, J. R. (Eds.). *Knowledge acquisition from text and pictures*. Amsterdam: Elsevier.
- Mayer, R. E. (1984). Aids to prose comprehension. *Educational Psychologist*, 19, 30–42.
- Mayer, R. E. (1989a). Models for understanding. *Review of Educational Research*, 59, 43–64.
- Mayer, R. E. (1989b). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology*, 81, 240–246.
- Mayer, R. E. (1992a). Guiding students' cognitive processing of scientific information in text. In M. Pressley, K. R. Harris, & J. T. Guthrie (Eds.), *Promoting academic competency and literacy in school* (pp. 243–258). San Diego: Academic.
- Mayer, R. E. (1992b). Knowledge and thought: Mental models that support scientific reasoning. In R. A. Duschl & R. J. Hamilton (Eds.), *Philosophy of science, cognitive psychology, and educational theory and practice* (pp. 226–243). Albany, NY: SUNY Press.
- Mayer, R. E. (1993a). Comprehension of graphics in text: An overview. *Learning and Instruction*, 3, 239–246.

- Mayer, R. E. (1993b). Illustrations that instruct. In R. Glaser (Ed.), *Advances in instructional psychology*, Vol. 5 (pp. 253–284). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Mayer, R. E. (1993c). Problem-solving principles. In M. Fleming & W. H. Levie (Eds.), *Instructional message design: Principles from the behavioral and cognitive sciences*. Englewood Cliffs, NJ: Educational Technology Publications.
- Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, 83, 484–490.
- Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84, 444–452.
- Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology*, 88, 64–73.
- Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, 82, 715–726.
- 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., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development*, 43(1), 31–44.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.
- Paas, F. G. W. C., & Van Merriënboer, J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, 6, 351–372.
- Reiber, L. P. (1990). Animation in computer-based instruction. *Educational Technology Research and Development*, 38, 77–86.
- Ross, S. M. (1994). Delivery trucks or groceries? More food for thought on whether media (will, may, can't) influence learning: Introduction to special issue. *Educational Technology Research and Development*, 42(2), 5–6.
- Salomon, G. (1979/1994). *Interaction of media, cognition, and learning*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Schnotz, W. (1993a). Introduction to special issue on comprehension of graphics in texts. *Learning and Instruction*, 3, 151–155.
- Schnotz, W. (1993b). On the relation between dual coding and mental models in graphics comprehension. *Learning and Instruction*, 3, 247–249.
- Schnotz, W., & Kulhavy, R. (Eds.). (1994). *Comprehension of graphics*. Oxford, England: Pergamon.
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge, England: Cambridge University Press.
- Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structure of technical material. *Journal of Experimental Psychology: General*, 119, 176–192.
- Wertheimer, M. (1959). *Productive thinking*. New York: Harper & Row.
- Wetzel, C. D., Radtke, P. H., & Stern, H. W. (1994). *Instructional effectiveness of video media*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Willows, D. M., & Houghton, H. A. (Eds.). (1987). *The psychology of illustration, Vol. 1: Basic research*. New York: Springer-Verlag.
- Wittrock, M. C. (1974). Learning as a generative activity. *Educational Psychologist*, 11, 87–95.
- Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist*, 24, 345–376.
- World Book Encyclopedia*. (1992). Chicago: World Book, Inc.
- Vosniadous, S., & Ortony, A. (1989). *Similarity and analogical reasoning*. Cambridge, England: Cambridge University Press.