MULTIMEDIA LEARNING

Richard E. Mayer

I. Introduction to Multimedia Learning

A. Purpose

For the last dozen years, my colleagues and I at the University of California, Santa Barbara (UCSB) have been investigating the nature of multimedia learning with a goal of building a research-based theory of how people learn from words and pictures. In this chapter, I summarize the fruits of this effort by presenting an introduction to key concepts in multimedia learning, a description of the materials we have used in our studies, a cognitive theory of multimedia learning, and a summary of nine theory-based effects that we have discovered in our research. In summarizing our nine theory-based effects, we draw on a corpus of 20 research articles that contain data for approximately 60 tests of our theory.

B. Definitions

Multimedia learning occurs when a learner builds a mental representation from words and pictures that have been presented. This definition is broad enough to include book-based environments consisting of text and illustrations, computer-based environments consisting of narration and animation, and virtual game environments consisting of interactive speech and animated microworlds.

For purposes of our research program, *multimedia instructional messages* (which we also refer to as *multimedia messages*) are presentations of material using words

0079-7421/02 $35.00

TABLE I

DEFINITIONS OF KEY TERMS

Term	Definition	Example
Multimedia learning	Building a mental representation from words and pictures	Building a mental model of the cause-and-effect system for lightning formation from a narrated animation (as summarized in Fig.1)
Multimedia instructional message	A presentation consisting of words and pictures that is intended to promote learning	A narrated animation on lightning formation (as summarized in Fig.1)

and pictures that are intended to foster learning. The words can be printed text (such as text printed in a window on a computer screen) or spoken text (such as speech presented via computer speakers). The pictures can be static graphics such as photos, drawings, maps, charts, figures, and tables or dynamic graphics such as video or animation. Table I summarizes the definitions and examples of the key terms: multimedia learning and multimedia instructional message.

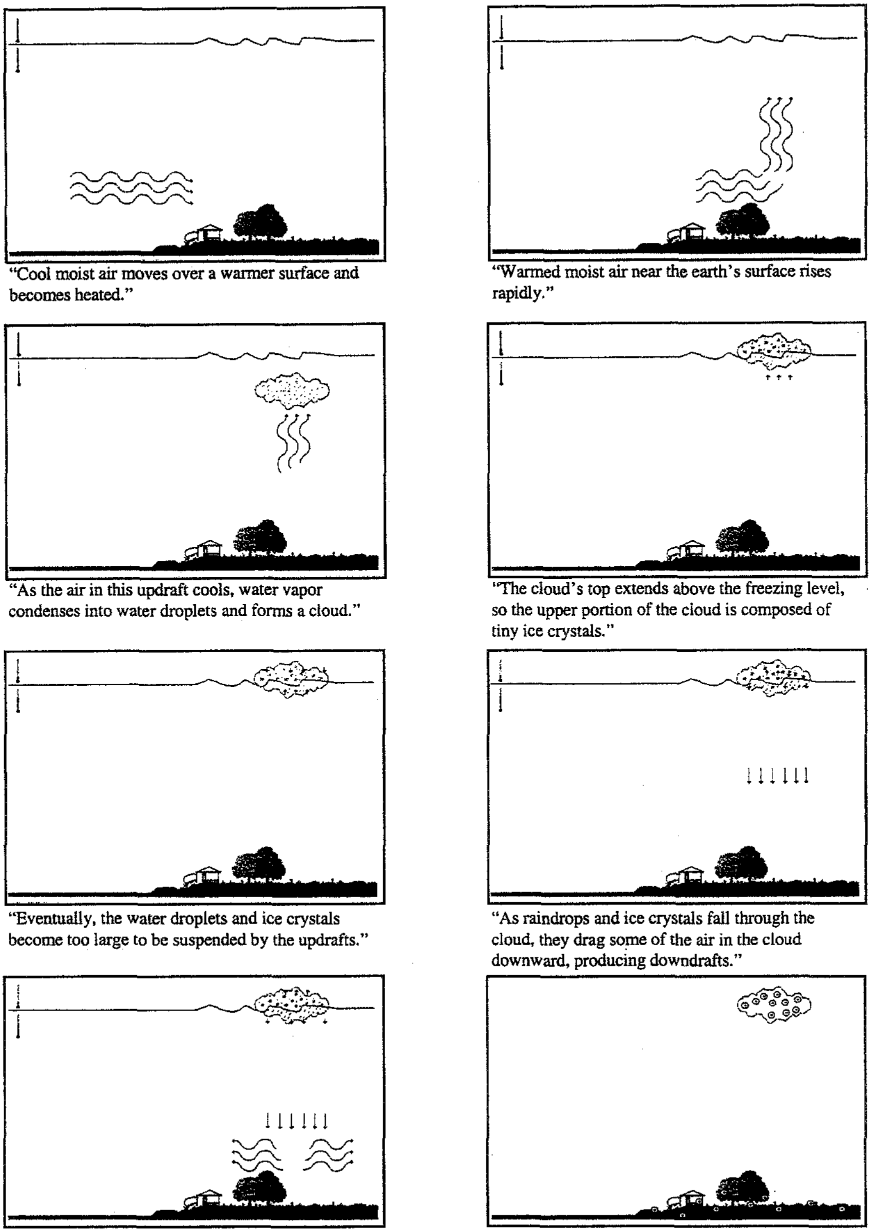
C. THE CASE FOR MULTIMEDIA LEARNING

Since its inception, the psychology of learning has favored verbal rather than pictorial forms of learning—dating back to the classic studies by Ebbinghaus (1885) on learning and remembering lists of nonsense syllables. When research methods shifted from classic verbal learning paradigms involving word lists and paired-associates, learning from prose became the dominant paradigm (Mayer, 1996, 2001a). However, with the advent of computer graphics and visualization tools, it is worthwhile to ask whether learning can be influenced when pictorial material is added to verbal material—a classic question that is being asked with increasing frequency (Mandl & Levin, 1989; Schnotz & Kulhavy, 1994; Willows & Houghton, 1987). In short, I am intrigued by the idea that human learners can achieve a deeper understanding when explanations are presented in the form of words and pictures rather than words alone. The goal of our research program is to explore the nature of multimedia learning—that is, learning from words and pictures.

D. EXAMPLES

1. Computer-Based Scenario

Consider the following scenario. You sit at a computer station, open a multimedia encyclopedia, and click on the entry for "lightning." As a result, you are presented with a 140-s narrated animation in which the narration describes the major steps in lightning formation and the corresponding animation depicts them. Figure 1



"Cool moist air moves over a warmer surface and becomes heated."

"Warmed moist air near the earth's surface rises rapidly."

"As the air in this updraft cools, water vapor condenses into water droplets and forms a cloud."

"The cloud's top extends above the freezing level, so the upper portion of the cloud is composed of tiny ice crystals."

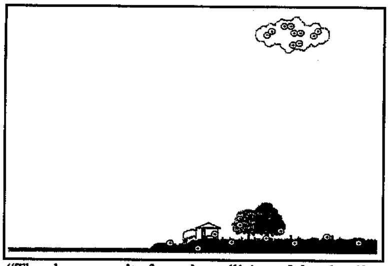
"Eventually, the water droplets and ice crystals become too large to be suspended by the updrafts."

"As raindrops and ice crystals fall through the cloud, they drag some of the air in the cloud downward, producing downdrafts."

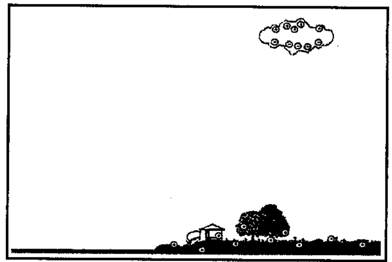
"When downdrafts strike the ground, they spread out in all directions, producing the gusts of cool wind people feel just before the start of the rain."

"Within the cloud, the rising and falling air currents cause electrical charges to build."

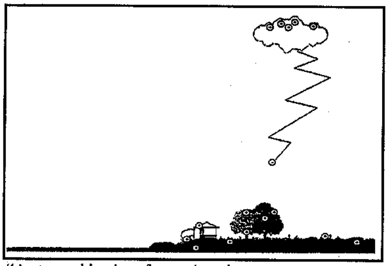
Fig. 1. Selected frames from a narrated animation on lightning formation. (From Mayer & Moreno, 1998. Copyright 1998 by the American Psychological Association. Reprinted by permission.)



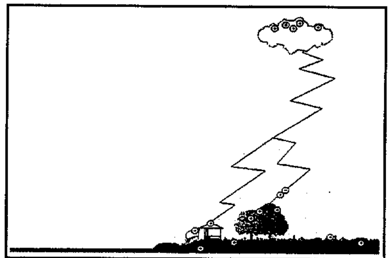
"The charge results from the collision of the cloud's rising water droplets against heavier, falling pieces of ice."



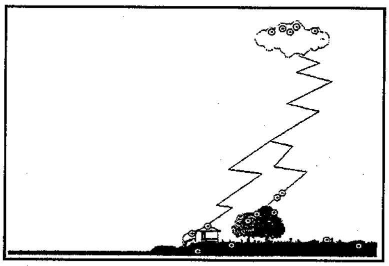
"The negatively charged particles fall to the bottom of the cloud, and most of the positively charged particles rise to the top."



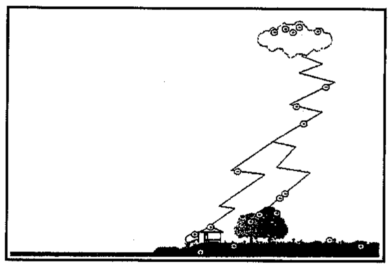
"A stepped leader of negative charges moves downward in a series of steps. It nears the ground."



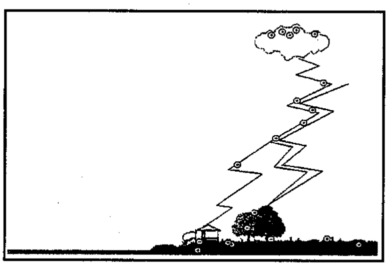
"A positively charged leader travels up from such objects as trees and buildings."



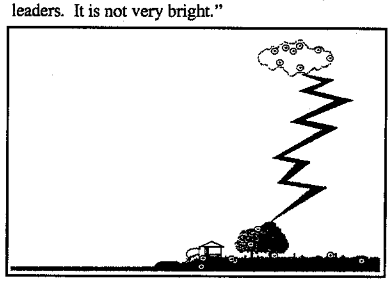
"The two leaders generally meet about 165-feet above the ground."



"Negatively charged particles then rush from the cloud to the ground along the path created by the leaders. It is not very bright."



"As the leader stroke nears the ground, it induces an opposite charge, so positively charged particles from the ground rush upward along the same path."



"This upward motion of the current is the return stroke. It produces the bright light that people notice as a flash of lightning."

Fig. 1. (*Continued*)

TABLE II

PROBLEM-SOLVING TRANSFER QUESTIONS FOR THE LIGHTNING LESSON

1. What could you do to decrease the intensity of lightning?
2. Suppose you see clouds in the sky but no lightning. Why not?
3. What does air temperature have to do with lightning?
4. What causes lightning?

presents selected frames from a narrated animation on lightning formation that we have used in some of our research studies (Mayer & Moreno, 1998). The lightning lesson is an example a *multimedia instructional message*—a presentation using words and pictures that is intended to foster learning. In this case, the words are represented as narration, the pictures are represented as animation, and the intended learning outcomes is a mental model of how the lightning system works.

Multimedia learning occurs if you construct a mental representation of the lightning system based on the words and pictures in the multimedia instructional message. In this case, you must build a cause-and-effect model of how a change in one part of the system causes a principle-based change in another part, and so on. For example, when cool air comes over a warm surface, the cool air becomes heated (by coming in contact with the warm surface) and rises (because hot air is less dense and less dense material rises).

To assess what you have learned, I ask you to answer a series of problem-solving transfer problems (each for 2.5 min) such as shown in Table II. In determining your transfer score, I give you one point for each acceptable answer on each question. For example, acceptable answers to the first question about decreasing the intensity of lightning include removing positively charged particles from the ground, removing negatively charged particles from the cloud, seeding the cloud with positively charged particles, placing an insulator between the cloud and the ground, and so on. I focus on tests of problem-solving transfer because I am interested in promoting meaningful learning, and transfer is a better measure of meaningful learning than is retention (Mayer, 2002).

As another example of a computer-based multimedia instructional message, consider a 45-s narrated animation about a car's braking system that is summarized in Fig. 2 (Mayer & Anderson, 1992). The narration describes the steps in the chain of activity when you press on the brake pedal, and the animation shows the steps in the chain. Table III lists some transfer questions for the learner to answer (each for 2.5 min). For example, some acceptable answers for the third question about "what could have gone wrong" include that the brake fluid leaked out of the tube, the piston is stuck in one position, or the brake shoe does not touch the drum.

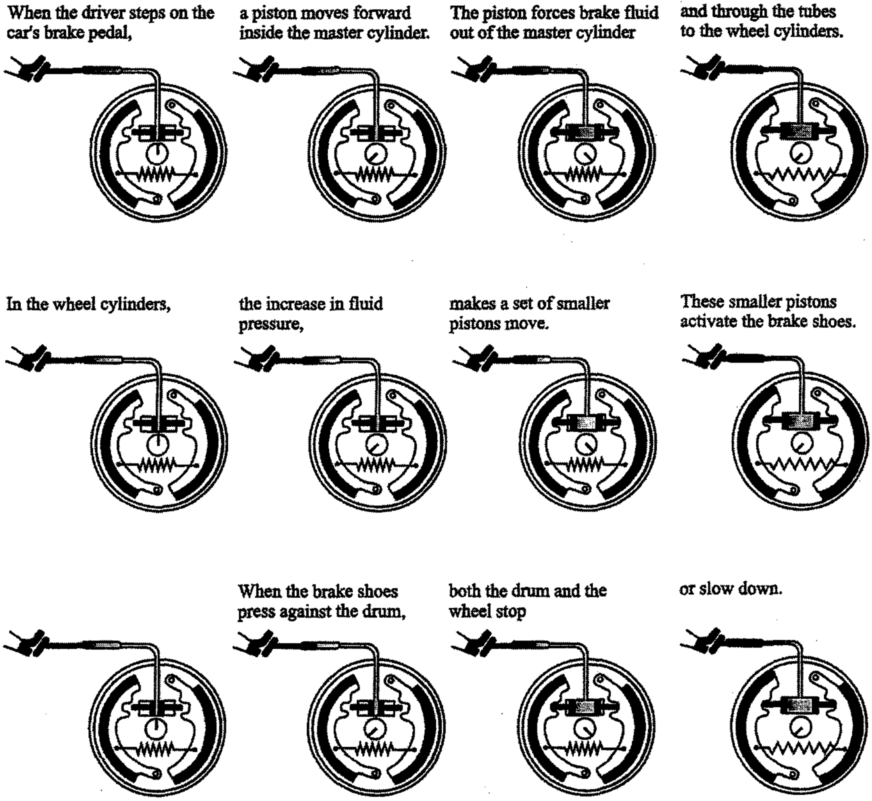


Fig. 2. Selected frames from a narrated animation on how brakes work. (From Mayer & Anderson, 1992. Copyright 1992 by the American Psychological Association. Reprinted by permission.)

TABLE III

Problem-Solving Transfer Questions for the Brakes Lesson

1. What could be done to make brakes more reliable, that is, to make sure they would not fail?
2. What could be done to make brakes more effective, that is, to reduce the distance needed to bring a car to a stop?
3. Suppose you press on the brake pedal in your car but the brakes don't work. What could have gone wrong?
4. What happens when you pump the brakes (i.e., press the pedal and release the pedal repeatedly and rapidly)?

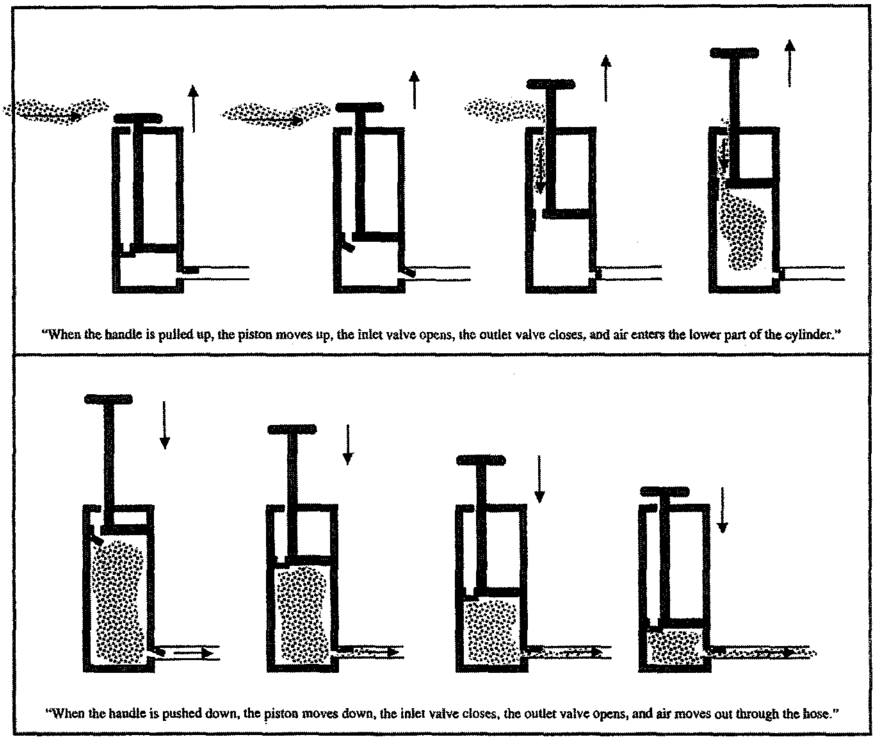


Fig. 3. Selected frames from a narrated animation on how a bicycle tire pump works. (Adapted from Mayer & Anderson, 1991.)

Finally, Fig. 3 shows a computer-based multimedia instructional message that explains how a bicycle tire pump works, and Table IV lists some transfer questions (Mayer & Anderson, 1991). For example, some acceptable answers for the second question about making the pump more effective include using a larger cylinder, a longer rod, or a tighter seal between the piston and cylinder. Other computer-based

TABLE IV

PROBLEM-SOLVING TRANSFER QUESTIONS FOR THE PUMP LESSON

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

materials we have used include a narrated animation explaining how the human respiratory system works (Mayer & Sims, 1994), how an electrical generator works (Mayer & Gallini, 1990), and how airplanes achieve lift (Mautone & Mayer, 2001).

2. *Book-Based Scenario*

Consider another scenario in which you are interested in a topic so you look it up in the pages of an encyclopedia. For example, Fig. 4 shows a book-based version of the lightning lesson based on five illustrations with captions (Mayer, Bove, Bryman, Mars, & Tapangco, 1996), Fig. 5 shows a book-based version of the brakes lesson (Mayer, 1989), and Fig. 6 shows a book-based version of the pump lesson (Mayer & Gallini, 1990). The same transfer questions can be given, as are shown in Tables II, III, and IV, respectively.

3. *Game-Based Scenario*

Finally, suppose you were playing an educational computer game designed to teach environmental science (Moreno & Mayer, 2000a; Moreno, Mayer, & Lester, 2000; Moreno, Mayer, Spires, & Lester, 2001). In the game—called Design-A-Plant—you travel to a distant planet that has certain environmental conditions such as heavy rainfall, lack of sun, and strong winds. With the assistance of an on-screen character named Herman the Bug you are asked to design a plant that would survive on the planet—by selecting appropriate roots, stem, and leaves. Along the way, Herman provides feedback in the form of narrated animation about plant growth. After traveling to several planets and designing plants for each one, you take a transfer test in which you must design a plant for a new planet and you must describe the environmental conditions best suited for a new plant. Your transfer score is based on the number of correct features in each of your answers.

In other studies we used a computer game aimed at teaching elementary school children how to add and subtract signed numbers in which the learner can move a bunny along a number line and receive multimedia feedback (Moreno & Mayer, 1999a).

Thus, our research on multimedia learning includes computer-based and book-based messages that explain how scientific systems work as well as educational games aimed at teaching scientific and mathematical concepts. In all cases, the multimedia instructional messages use words and pictures to help people learn.

E. Three Views of Multimedia Instructional Messages

Mayer (2001b) has examined three views of multimedia instructional messages: the delivery media view (which is based on devices used to deliver the message),

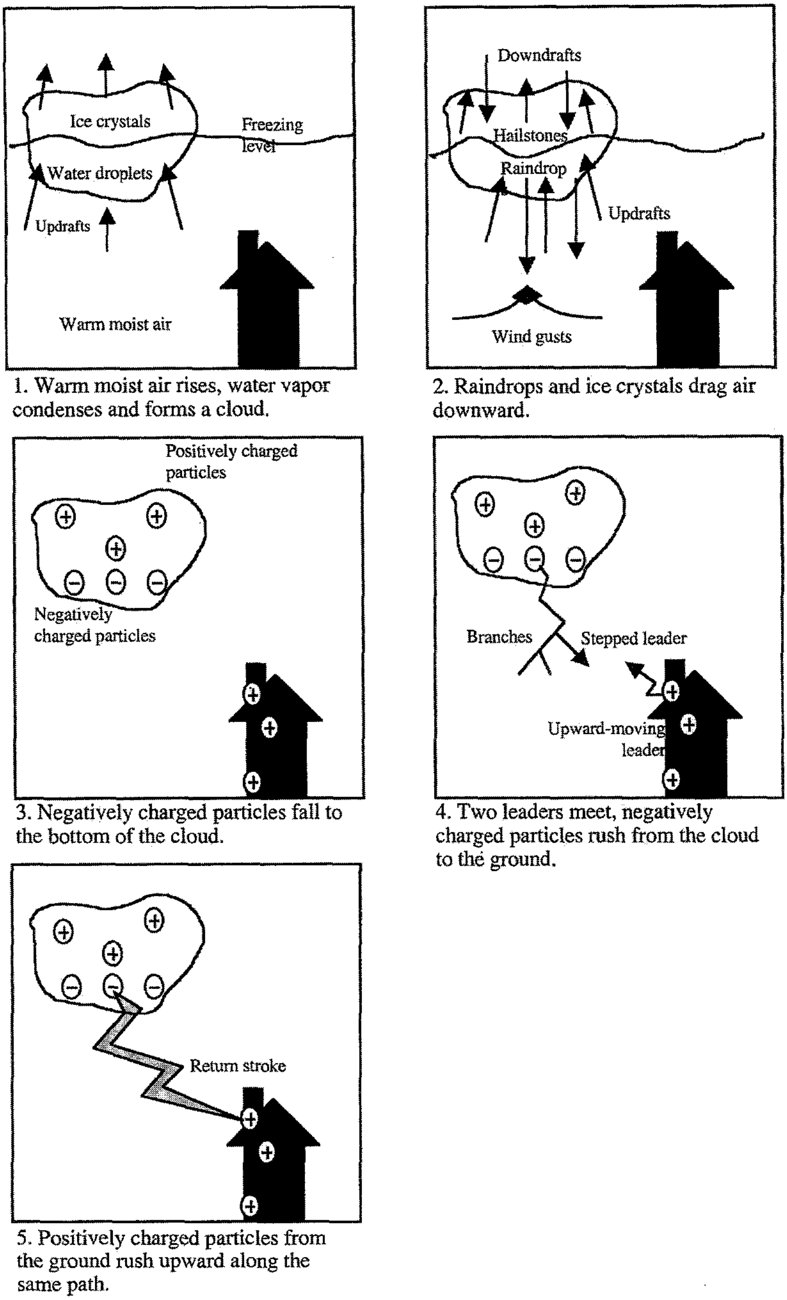


Fig. 4. Annotated illustrations for a book-based lesson on lightning formation. (From Mayer et al., 1996. Copyright 1996 by the American Psychological Association. Reprinted by permission.)

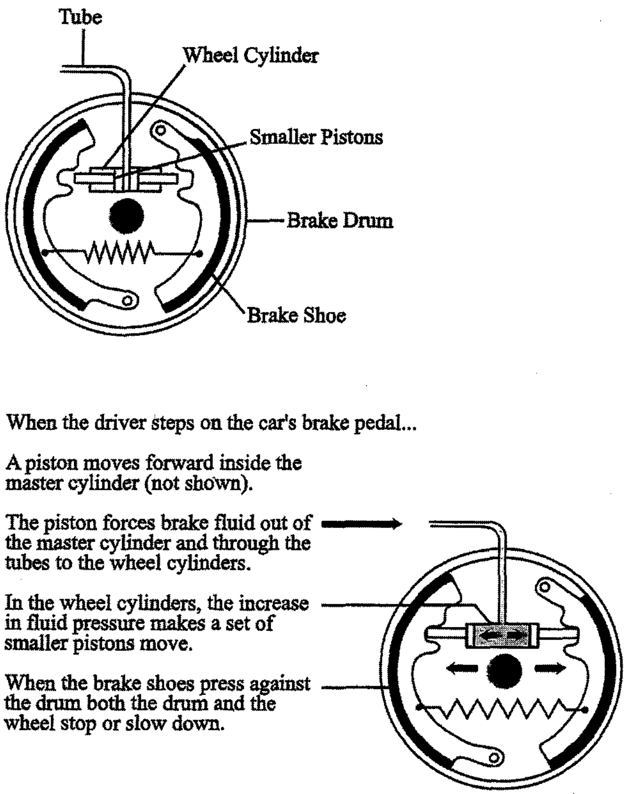


Fig. 5. Annotated illustrations for a book-based lesson on how brakes work. (Adapted from Mayer, 1989.)

the presentation modes view (which is based on the representational formats used to present the material), and the sensory modalities view (which is based on the sense modalities the learner uses to receive the message). The three views are summarized in Table V.

According to the delivery media view, multimedia messages involve two or more delivery devices such as paper, human voice boxes, blackboards, computer screens, amplified speakers, headphones, head-mounted displays, tape players, CD players, overhead projectors, and VCRs. For example, in a book-based environment, ink on paper appears to be the only presentation device (unless you wish to define text portions of the page and graphics portions of the page as different delivery devices); in a classroom environment, the delivery devices could be the instructor's voice and the projection system for PowerPoint slides; in a computer-based environment, we can present material both on a computer screen

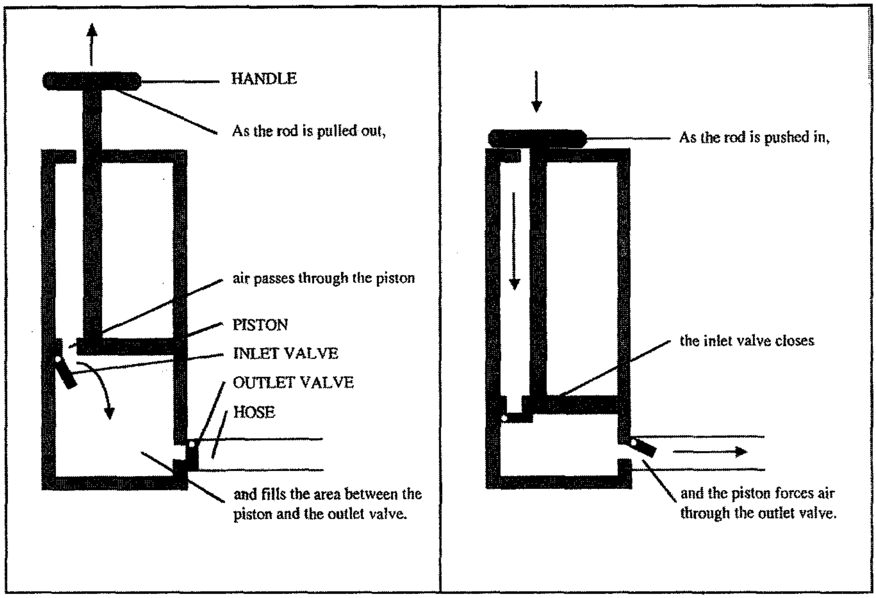


Fig. 6. Annotated illustrations for a book-based lesson on how a bicycle tire pump works. (Adapted from Mayer & Gallini, 1990.)

and through amplified speakers; in a virtual game environment, we can present material both through the head-mounted display and via headphones. Thus, in the delivery media view, a multimedia message exists when two or more devices are used to deliver the information. Accordingly, the narrated animation summarized in Fig. 1 is a multimedia message because it is delivered via computer screen and amplified speakers.

TABLE V

THREE VIEWS OF MULTIMEDIA

View	Definition	Example of narrated animation (in Fig. 1)
Delivery media	Two (or more) delivery devices	Amplified speakers (narration) and computer screen (animation)
Presentation modes	Verbal and pictorial representations	Words (narration) and pictures (animation)
Sensory modalities	Auditory and visual senses	Sounds (narration) and pictures (animation)

Adapted from Mayer (2001b). Reprinted by permission.

Although the delivery media view is quite objective in determining which devices are being used to present a message, it does not offer much help in building a psychological account of how people learn. Overall, I reject the delivery media view because it focuses on technology rather than learners.

According to the presentation modes view, multimedia messages involve verbal and pictorial forms of representation (i.e., words and pictures), such as animation with narration, animation with on-screen text, diagrams with narration, or diagrams with printed text. In each case, the presentation includes words (in the form of narration or printed text) and pictures (in the form of animation or diagrams). Thus, the narrated animation depicted in Fig. 1 is a multimedia message because it consists of words (i.e., the narration) and pictures (i.e., the animation).

According to the sensory modalities view, multimedia messages involve auditory and visual modalities of representation, such as animation with narration or diagrams with narration. In each case, the presentation involves auditory material (in the form of narration) and visual material (in the form of animation or diagrams). Printed words might initially be processed in visual form but the learner may mentally convert the printed words to sounds, thereby creating auditory representations. Thus, the narrated animation depicted in Fig. 1 is a multimedia message because it consists of auditory material (i.e., the narration) and visual material (i.e., the animation).

Both the presentation modes view and the sensory modalities view offer a psychological basis for defining multimedia—based on the representation code (i.e., verbal or pictorial) used by the learner (i.e., presentation modes view) or based on the representation modality (i.e., auditory or visual) used by the learner (i.e., sensory modalities). Similarly, Paivio's (1986) dual code theory relies on a presentation modes view, whereas Baddeley's (1992, 1999) working memory theory relies on a sensory modalities view. I use elements of both views in the cognitive theory of multimedia learning, but future research is needed to clarify the respective merits of each view and to reach some sort of reconciliation (Miyake & Shah, 1999). In defining multimedia messages as containing words and pictures, I am most closely following the presentation modes approach, but this definition is also consistent with the sensory modalities view if you assume that the learner mentally converts printed words into sounds.

F. Two Approaches to the Design of Multimedia Instructional Messages

Building on analyses by Norman (1993), I (Mayer, 1999a, 2001b) have distinguished between two approaches to the design of multimedia instructional messages—a *technology-centered approach* and a *learner-centered approach.* The features of the two approaches are summarized in Table VI.

In a technology-centered approach, one focuses on the nature of the delivery technology as the starting point for message design. In general, designers who

TABLE VI

TWO APPROACHES TO MULTIMEDIA DESIGN

Design approach	Starting point	Goal	Issues
Technology-centered	Capabilities of multimedia technology	Provide access to information	How to use cutting-edge technology in designing multimedia messages
Learner-centered	Workings of the human mind	Aid human cognition	How to adapt multimedia technology to support human cognitive processing

Adapted from Mayer (2001b). Reprinted by permission. Copyright 2001 Cambridge University Press.

take a technology-centered approach focus on a cutting edge technology, and ask: "How can I use this technology to deliver multimedia instructional messages?" For example, beginning in the 1990s, technology-centered designers have been seeking ways to use the Internet as a delivery system for multimedia instruction.

My problem with the technology-centered approach is that it ignores the nature of the learner. It does not take into account an understanding of how people learn. In short, it requires that the learner adapt to the new technology.

In a learner-centered approach, one focuses on the nature of human learning as the starting point for message design. In general, designers who take a learner-centered approach focus on a theory of how people learn and ask: "How can I use this technology to support and enhance human learning?" For example, beginning in the 1990s, learner-centered designers have been seeking ways to use the Web-based multimedia as an tool for improving human learning. In the learner-centered approach, the technology must be adjusted to fit the needs of the learner. In taking a learner-centered approach, I begin with the premise that principles of multimedia design should be based on a theory of how people learn (Bransford, Brown, & Cocking, 1999; Lambert & McCombs, 1998).

If you examine the history of technology-centered approaches to education, you are likely to find many disappointments. Cuban (1986) has shown how implementations of educational technology during the 20th century typically followed a cycle of strong claims, followed by large-scale implementation in schools, followed by massive disappointment. For example, in the 1920s, motion pictures were seen as the hot new technology that would revolutionize education. Thomas Edison in 1922 predicted that "the motion picture is destined to revolutionize our educational system and that in a few years it will supplant largely, if not entirely, the use of textbooks" (cited in Cuban, 1986, p. 9). Foreshadowing current claims about the potential of computer graphics, proponents of motion pictures believed that educational concepts could be best communicated through dynamic visual presentations. In spite of such claims and many attempts to implement the technology of motion pictures, reviews of educational practice show that teachers rarely use educational films in their classrooms (Cuban, 1986).

During the 1930s and 1940s, attention shifted to the educational potential of radio. Reminiscent of current claims for the World Wide Web, Benjamin Darrow in 1932 proposed that radio would "bring the world to the classroom, to make universally available the services of the finest teachers, the inspiration of the greatest leaders. . . ." As school systems rush to get every classroom connected to the Web, it is humbling to acknowledge that "radio has not been accepted as a full-fledged member of the educational community" (Cuban, 1986, p. 24).

During the 1950s, educational television was touted as a miracle technology that combined the visual power of motion pictures with the worldwide coverage of radio. Postwar schools would become "continental classrooms" that provided "richer education at less cost" (Cuban, 1986, p. 33). Yet, reviews of educational practice show that teachers rarely use educational television in their classrooms (Cuban, 1986).

During the 1960s and 1970s, computer-based programmed instruction was offered as a technology that would solve the problems of education. The optimistic view of computer-based instruction lead to large-scale implementations in the 1970s, such as PLATO and TICCIT, but subsequent research showed that such systems did not necessarily lead to better learning than did conventional teacher-led instruction (Cognition and Technology Group at Vanderbilt, 1996).

What lesson is to be learned from the history of educational technology in the 20th century? Although film, radio, television, and programmed instruction each relied on the cutting-edge technologies of their day, they all failed to seriously affect student learning and educational practice. In each case, a technology-centered approach was used in which reformers asked learners to adjust to the requirements of current educational technologies. Claims for the educational potential of visual learning and worldwide learning have resurfaced in the context of Web-based multimedia learning. Will the latest educational technology meet the same fate as its predecessors? In my opinion, that depends on whether designers take a technology-centered approach or a learner-centered approach.

G. Two Metaphors of Multimedia Learning

The way one designs a multimedia message depends on one's conception of how people learn. In this section, I distinguish between two metaphors of learning that have evolved during psychology's first century of scientific study (Mayer, 1992, 2001a)—*multimedia learning as information acquisition* and *multimedia learning as knowledge construction*. The features of these metaphors of multimedia learning are summarized in Table VII.

The most straightforward view is that multimedia learning occurs when a learner adds new information to memory. The instructor's job is to present information to the learner, and the learner's job is to receive the information. The information

TABLE VII

TWO METAPHORS OF MULTIMEDIA LEARNING

Metaphor	Definition	Content	Learner	Teacher	Goal
Information acquisition	Adding information to memory	Information	Passive information recipient	Information provider	Deliver information; be a delivery vehicle
Knowledge construction	Building a coherent cognitive structure	Knowledge	Active sense maker	Cognitive guide	Provide cognitive guidance; be a helpful communicator

acquisition view can also be called the transmission view because information is transmitted from instructor to learner. The information acquisition view grows out of information-processing theories of learning that were popular in the 1970s (Lachman, Lachman, & Butterfield, 1979). According to the information acquisition view, multimedia environments can be seen as delivery systems because their role is to deliver words and pictures to the learner as efficiently as possible.

What is wrong with the information acquisition metaphor of multimedia learning? The major problem is that it is offers a limited vision of learning that is not consistent with modern research on how people learn (Bransford et al., 1999; Lambert & McCombs, 1998). When people engage in deep learning, they work actively to make sense of the presented material—an approach that is inconsistent with the passive learner portrayed in the information acquisition view.

An important alternative to the information acquisition view is the knowledge construction view of multimedia learning. Although the knowledge construction view has its roots in classic work on meaningful learning by Bartlett (1932) and Piaget (1954), the modern rebirth of the knowledge construction view began in the 1980s as a reaction against the limitations of the information acquisition view (Mayer, 1996, 2001b). According to the knowledge construction view, learning occurs when a learner engages in active cognitive processing with a goal of making sense of incoming material. The instructor is a cognitive guide who helps the learner engage in appropriate cognitive processing, and the learner is an active cognitive processor attempting to make sense of the presented material. According to the knowledge construction view, multimedia systems are not delivery devices but rather are venues for fostering the process of sense making by learners. This is the metaphor of multimedia learning that drives our research program.

H. Two Kinds of Multimedia Learning Outcomes

There are two methods for evaluating what someone has learned—*retention tests* and *transfer tests* (Anderson et al., 2001; Mayer & Wittrock, 1996). Retention tests measure how much someone remembers and can be in the form of recall problems (e.g., "Please write down everything you can remember from the lesson you just received."), or recognition problems (e.g., a multiple-choice test or a true–false test). Transfer tests measure how well someone can apply what they learned to a new situation and can be measured in the form of essay items (e.g., "Please explain what would happen if...") or multiple-choice items (e.g., "Which of the following events would happen if..."). I prefer to focus on transfer tests because they provide the best measure of learner understanding.

It is customary to distinguish between *rote* and *meaningful learning outcomes* (Anderson et al., 2001; Mayer, 2002), as summarized in Table VIII. Rote learning outcomes consist of pieces of information represented in memory in much the same way as they were presented. Memorizing the verbatim definition of a

TABLE VIII

Two Kinds of Multimedia Learning Outcomes

Learning outcome	Cognitive description	Test performance	
		Retention	Transfer
Rote learning	Fragmented, isolated knowledge	Good	Poor
Meaningful learning	Organized, integrated knowledge	Good	Good

term is an example of a rote-learning outcome. In our research, we define rote learning as indicated by good performance on retention and poor performance on transfer.

Meaningful learning outcomes consist of coherent mental representations (such as mental models) that are internally organized and externally connected to other knowledge. Understanding how a device works is an example of meaningful learning. In our research, we define meaningful learning as indicated by good retention performance and good transfer performance.

I. Two Kinds of Active Learning

Active learning can be defined in terms of behavioral activity or cognitive activity, as summarized in Table IX. Behaviorally active learning occurs when the learner is physically engaged in learning activities—such as pressing buttons on a keyboard, moving a joystick, or using a mouse to click on words or pictures on a computer screen. The epitome of behaviorally active learning is hands-on activity (or a computer-based simulation of hands-on activity), such as conducting a science experiment. Behaviorally passive learning occurs when the learner is not physically engaged in learning activities; for example, the epitome of behaviorally passive learning is sitting quietly as information is presented (either in a lecture or via a computer-based multimedia presentation).

In contrast, cognitively active learning occurs when the learner engages in deep cognitive processing during learning such as paying attention to relevant material, mentally organizing the material into a coherent structure, and mentally integrating the structure with existing knowledge. The epitome of cognitively active learning

TABLE IX

Distinction between Cognitive Activity and Behavioral Activity

	Low cognitive activity	High cognitive activity
Low behavioral activity	Does not foster meaningful learning	Does foster meaningful learning
High behavioral activity	Does not foster meaningful learning	Does foster meaningful learning

is the construction of a *mental model*—that is, a coherent mental representation of a system in which the parts can interact with one another in a principled way. Cognitively passive learning occurs when the learner does not engage in active cognitive processing during learning; for example, the epitome of cognitively passive learning is rote learning of incoming material.

A guiding premise of our research is that meaningful learning is caused by cognitive activity during learning rather than behavioral activity during learning, whereas rote learning is caused by low levels of cognitive activity during learning. As shown in Table IX, cognitively active learning can occur with behaviorally active learning or with behaviorally passive learning, and cognitively passive learning can occur with behaviorally active learning or with behaviorally passive learning. Thus, our research on multimedia learning focuses on the fostering of cognitive activity during learning.

II. A Cognitive Theory of Multimedia Learning

In order to understand how to foster multimedia learning, it is useful to begin with a research-based theory of how people learn from words and pictures. A major goal of our research program is to test a cognitive theory of multimedia learning that can account for how people learn from words and pictures in multimedia environments. To construct a cognitive theory of multimedia learning, we draw on three assumptions derived from current research in cognitive science: the dual-channel assumption, the limited capacity assumption, and the active processing assumption. These assumptions are summarized in Table X.

The dual-channel assumption is that humans possess separate information processing systems for visual (or pictorial) material and auditory (or verbal) material. There are two ways of characterizing the channels—by sensory modality (such as visual versus auditory) as proposed by Baddeley (1992, 1999) or by presentation modes (such as pictorial versus verbal) as proposed by Paivio (1986). For purposes of our theory, we have opted for a compromise in which early processing is based on sense modality and late processing is based on presentation mode. For example, on-screen text is initially processed in an visual/pictorial channel because it enters the cognitive system through the eyes, and subsequently is converted mentally into sounds for processing in the auditory/verbal channel.

The limited capacity assumption is that humans are limited in the amount of material they can process in each channel at one time. Capacity limitations are a central feature of Chandler and Sweller's (1991; Sweller, 1999) cognitive load theory and Baddeley's (1992, 1999) model of working memory, as well as other models of working memory (Miyake & Shah, 1999). For purposes of our theory, we assume that each channel has a limited capacity such that presenting too much material on the screen at one time can overload the visual/pictorial channel

TABLE X

THREE ASSUMPTIONS OF A COGNITIVE THEORY OF MULTIMEDIA LEARNING

Assumption	Description	Related citations
Dual channels	Humans possess separate channels for processing visual and auditory information	Paivio, 1986; Baddeley, 1992
Limited capacity	Humans are limited in the amount of material they can process in each channel at one time	Baddeley, 1992; Chandler & Sweller, 1991
Active processing	Humans engage in active learning by attending to relevant incoming material, organizing selected material into a coherent mental representation, and integrating mental representations with other knowledge	Mayer 1999b, Wittrock, 1989

Adapted from Mayer (2001b). Reprinted by permission. Copyright 2001 Cambridge University Press.

and presenting too much material via the speakers at one time can overload the auditory/verbal channel.

The active processing assumption is that humans engage in meaningful learning by attending to relevant incoming information, organizing selected material into coherent mental representations, and integrating mental representations with other knowledge. This assumption follows from Wittrock's (1989) generative learning theory and Ausubel's (1968) assimilative learning theory. The three cognitive processes correspond to Mayer's (1996, 1999b) selecting, organizing, and integrating, and Kintsch's (1998) view of prose processing as building surface, textbase, and situation model representations. For purposes of our theory, we assume that within each channel, and subject to capacity limitations, active learners select, organize, and integrate aspects of incoming words and pictures.

Figure 7 represents the cognitive theory of multimedia learning as a series of boxes and arrows. The top row represents the auditory/verbal channel, whereas

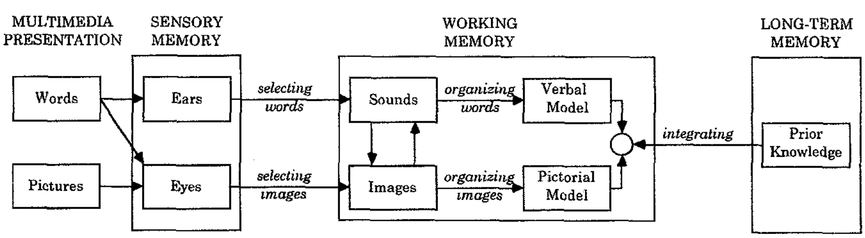


Fig. 7. A cognitive theory of multimedia learning. (Adapted from Mayer 2001. Reprinted by permission. Copyright 2001 Cambridge University Press.)

the bottom row represents the visual/pictorial channel. The first column represents the material in a multimedia presentation—namely words or pictures. The second column represents the learner's sensory memory—namely sensory images from the ears or eyes. The third and fourth columns represent processing in working memory—first, the mental representation of sounds and visual images and later, the mental representation of verbal and pictorial models. The fifth column represents the learner's storehouse of prior knowledge in long-term memory.

On the top row, the arrow from words to ears represents sound waves from spoken words impinging on the learner's ears; the arrow from ears to sounds (labeled *selecting words*) represents paying attention to some of the incoming spoken words; and the arrow from sounds to verbal model (labeled *organizing words*) represents mentally organizing the selected words into a coherent representation. On the bottom row, the arrow from pictures to eyes represents light waves from graphics impinging on the learner's eyes; the arrow from words to eyes represents light waves from printed words impinging on the learner's eyes; the arrow from eyes to images (labeled as *selecting images*) represents paying attention to some of the incoming visual material; and the arrow from images to pictorial model (labeled *organizing images*) represents mentally organizing the selected images into a coherent mental representation. Finally, the arrows from verbal model, pictorial model, and prior knowledge (labeled *integrating*) represent mentally combining the verbal and pictorial models with each other and with relevant prior knowledge in the learner's long-term memory.

Consider the course of information processing when a narrated animation is presented (such as is summarized in Fig. 1). In the auditory/verbal channel, words are presented as narration, which are detected by the ears; the learner pays attention to some of the incoming sounds (indicated by the selecting words arrow) resulting in some of the words being represented in working memory. Then, the learner mentally organizes the words into a verbal model (indicated by the organizing words arrow). Meanwhile, back in the visual/pictorial channel, pictures are presented as animation frames which are detected by the eyes; the learner pays attention to some aspects of the incoming images (indicated by the selecting images arrow) resulting in some of the images being represented in working memory. Then, the learner mentally organizes the visual images into a pictorial model (indicated by organizing images arrow). Finally, the learner mentally integrates the verbal model and pictorial model with each other and with relevant knowledge from long-term memory (indicated by the integrating arrows), resulting in a meaningful learning outcome that can be stored in long-term memory.

In summary, in constructing the cognitive theory of multimedia learning I adapted three basic themes in cognitive science—dual channels, limited capacity, and active processing—to a multimedia-learning environment. Based on this model, I can predict that multimedia messages that foster all five of the cognitive processes shown in Fig. 7 are more likely to lead to meaningful learning than are those that do not foster all of the cognitive processes. In particular, for meaningful learning to occur, multimedia messages must be constructed to enable the learner to hold corresponding verbal and pictorial material in working memory at the same time. In the following sections, I explore several of the key predictions of the cognitive theory of multimedia learning along with reviews of research testing the predictions.

III. Multimedia Effect

A. What Is the Multimedia Effect?

People can learn more deeply when they receive an explanation in words and pictures rather than words alone. This assertion is fundamental to the case for multimedia learning, so it is useful to examine whether it has any empirical support. In this section, I compare the transfer test performance of students who receive an explanation in words with the transfer test performance of students who receive an explanation in words and pictures. For example, we can present the verbal narration of how lightning storms develop, as listed in the frames of Fig. 1 (words-alone presentation) or we can present the same narration along with corresponding animation depicting the events in the narration (words-and-pictures presentation). A *multimedia effect* (for transfer) occurs if learners perform better on transfer tests when they receive an explanation in words and pictures rather than in words alone.

B. Theory: What Are the Mechanisms Underlying the Multimedia Effect?

The information delivery theory holds that multimedia messages are vehicles for delivering information to the learner. According to the information delivery theory, words and pictures are informationally equivalent—that is, the narration describing the steps in lightning formation contains the same information as the animation depicting the steps in lightning formation. Therefore, words-alone presentations should result in equivalent learning outcomes as do words-and-pictures presentations, because the same information is delivered in both presentations.

In contrast, the cognitive theory of multimedia learning holds that deep learning occurs when learners engage in all five cognitive processes listed in Fig. 7—selecting words, selecting images, organizing words, organizing images, and integrating. The processes involving images are less likely to occur with words-only presentations than with words-and-pictures presentations. Therefore, words-alone presentations should result in poorer transfer performance than words-and-pictures presentations.

The case for presenting words and pictures rather than words alone is consistent with aspects of Paivio's dual coding theory (1986), particularly the idea that better learning occurs when learners build referential connections between verbal and nonverbal mental representations of the same item. Memory for a presented item is maximized when learners build verbal and nonverbal representations of the item, and build referential connections between the representations. Although dual coding theory was not developed to account for meaningful learning from multimedia messages, it has been extended recently to account for text reading and writing (Sadoski & Paivio, 2001) and is incorporated into the cognitive theory of multimedia learning (Mayer, 2001b).

C. Research: Is There a Multimedia Effect?

1. *Core Findings*

Do students understand an explanation more deeply from words and pictures than from words alone? To address this question, I identified 11 experiments in which my colleagues and I compared the transfer performance of students who received a words-alone presentation and those who had received a words-and-pictures presentation. The comparisons included computer-based explanations of how pumps work (Mayer & Anderson, 1991, Experiment 2a; Mayer & Anderson, 1992, Experiment 1), how brakes work (Mayer & Anderson, 1992, Experiment 2), and how lightning storms develop (Moreno & Mayer, in press, Experiment 1) in which we compared narration alone to narration and animation; paper-based explanations of how brakes work (Mayer, 1989, Experiments 1 and 2; Mayer & Gallini, 1990, Experiment 1), how pumps work (Mayer & Gallini, 1990, Experiment 2), how electrical generators work (Mayer & Gallini, 1990, Experiment 3), and how lightning storms develop (Mayer et al., 1996, Experiment 2) in which we compared printed text to printed text and illustrations; and game-based explanations of how to add and subtract signed numbers (Moreno & Mayer, 1999, Experiment 1) in which we compared showing answers in terms of arithmetic symbols to showing answers in terms of arithmetic symbols along with animation. Table XI shows the effect sizes on transfer tests comparing words-alone versus words-and-pictures groups across the 11 studies. In 11 out of 11 comparisons, the learners who received words and pictures performed better on transfer tests than did students who received words alone, yielding a median effect size of 1.39.

These results provide clear and consistent support for a *multimedia effect:* People learn more deeply from words and pictures than from words alone. The multimedia effect is consistent with the predictions of the cognitive theory of multimedia learning and inconsistent with the predictions of the information delivery theory.

TABLE XI

SUMMARY OF MULTIMEDIA EFFECTS: BETTER TRANSFER WHEN A MESSAGE CONTAINS WORDS AND PICTURES RATHER THAN WORDS ALONE

Source	Medium	Content	Effect size
Mayer & Anderson, 1991, Exp. 2a	Screen	Pumps	2.43
Mayer & Anderson, 1992, Exp. 1	Screen	Pumps	1.90
Mayer & Anderson, 1992, Exp. 2	Screen	Brakes	1.67
Moreno & Mayer, 2001, Exp. 1	Screen	Lightning	0.50
Mayer, 1989, Exp. 1	Page	Brakes	1.50
Mayer, 1989, Exp. 2	Page	Brakes	1.71
Mayer & Gallini, 1990, Exp. 1	Page	Brakes	1.19
Mayer & Gallini, 1990, Exp. 2	Page	Pumps	1.00
Mayer & Gallini, 1990, Exp. 3	Page	Generators	1.35
Mayer et al., 1996, Exp. 2	Page	Lightning	1.39
Moreno & Mayer, 1999, Exp. 1	Game	Arithmetic	0.47
Median			1.39

2. *Related Research*

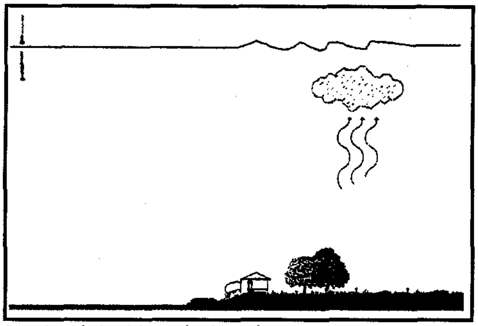
The graphics we used in our studies were intended to explain how something works—and could be called *explanative*. In contrast, *nonexplanative* graphics may be used to decorate a page (such as a picture of a car driving down the road in a passage about car brakes) or to portray a single element (such as a picture of a lightning bolt under the heading "lightning bolt"). In a survey of science textbooks (Mayer, 1993) and mathematics textbooks (Mayer, Sims, & Tajika, 1995), we found that the overwhelming majority of illustrations were nonexplanative. It is unlikely that nonexplanative graphics would produce a multimedia effect, although this prediction has not been subjected to intensive testing.

IV. Spatial Contiguity Effect

A. WHAT IS THE SPATIAL CONTIGUITY EFFECT?

The foregoing section provided evidence that students can understand more deeply from words and pictures than from words alone. However, all multimedia messages (i.e., presentations with words and pictures) are not equally effective. A major focus of our research program—as reflected in the remainder of this chapter—is to determine the conditions under which multimedia presentations promote deep understanding. I begin with an examination of the placement of corresponding pictures and printed words on a page or screen.

Separated Presentation



As the air in this updraft cools, water vapor condenses into water droplets and forms a cloud.

Integrated Presentation

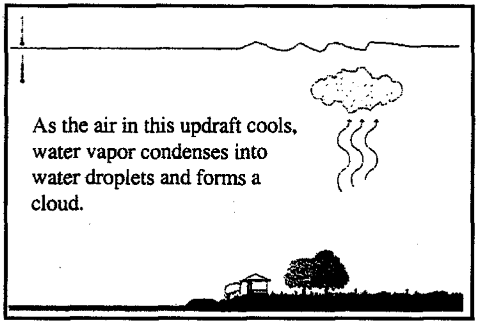


Fig. 8. Selected frames from an integrated and separated presentation of the lightning lesson using animation and on-screen text. (From Moreno & Mayer, 1999b. Copyright 1999 by the American Psychological Association. Reprinted by permission.)

For example, Fig. 8 shows frames from a computer-based presentation on lightning formation involving animation and on-screen text in which each sentence of on-screen text is placed next to the event it describes (integrated presentation) or at the bottom of the screen (separated presentation). Similarly, in book-based presentations on lightning formation, we can present a series of illustrations depicting the steps in lightning formation along with embedded captions that describe the steps in words (integrated presentation) or we can present the illustrations on one page and the printed words on another page (separated presentation). A spatial contiguity effect occurs if students perform better on transfer tests after receiving the integrated presentation rather than the separated presentation.

B. THEORY: WHAT ARE THE MECHANISMS UNDERLYING THE SPATIAL CONTIGUITY EFFECT?

According to a strict interpretation of the information delivery theory, identical information is presented with integrated and separated presentation. Thus, students in the integrated and separated groups should perform about the same on transfer tests. According to a lenient interpretation of the information delivery theory, separated presentation is like getting two separate deliveries of the same information, whereas integrated presentation is like getting just one delivery of the information. Thus, students in the separated group get essentially twice as much exposure to the information as students in the integrated group, and therefore should perform better on tests of transfer.

In contrast, the cognitive theory of multimedia learning posits that students are more likely to be able to hold corresponding verbal and pictorial representations in working memory at the same time with integrated presentation rather than separated presentation. Students in the separated group must expend their limited cognitive resources to search for pictures that correspond to text segments (or vice versa), whereas the placement of the words and pictures in the integrated group enables students to use their limited cognitive resources to engage in deep cognitive processing. Thus, students in the integrated group should perform better on transfer tests than students in the separated group.

C. RESEARCH: IS THERE A SPATIAL CONTIGUITY EFFECT?

1. Core Findings

Table XII summarizes the results of five comparisons between integrated and separated presentation of printed words and illustrations in a book-based venue involving lightning (Mayer, 1989, Experiment 2) and brakes (Mayer, Steinhoff,

TABLE XII

SUMMARY OF SPATIAL CONTIGUITY EFFECTS: BETTER TRANSFER WHEN PRINTED WORDS ARE PLACED NEAR RATHER THAN FAR FROM CORRESPONDING PICTURES

Source	Medium	Content	Effect size
Mayer, 1989, Exp. 2	Page	Brakes	1.71
Mayer et al., 1995, Exp. 1	Page	Lightning	1.09
Mayer et al., 1995, Exp. 2	Page	Lightning	1.35
Mayer et al., 1995, Exp. 3	Page	Lightning	1.12
Moreno & Mayer, 1999, Exp. 1	Screen	Lightning	0.48
Median			1.12

Bower, & Mars, 1995, Experiments 1, 2, and 3) and in a computer-based venue involving lightning (Moreno & Mayer, 1999b, Experiment 1). As you can see, in five out of five comparisons, integrated presentation lead to better transfer performance than did separated presentation, yielding a median effect size of 1.12. These results are consistent with the cognitive theory of multimedia learning and inconsistent with the information delivery theory. We refer to this pattern of results as a *spatial contiguity effect:* Students learn more deeply when corresponding pictures and printed words are placed near rather than far from each other on the page or screen.

2. *Related Research*

Sweller and his colleagues (Chandler & Sweller, 1992; Sweller, 1999) use the term *split attention effect* to refer to any impairment in learning caused by students having to integrate disparate sources of information. Consistent with the cognitive theory of multimedia learning, when students must use cognitive resources to search across disparate sources, they have fewer cognitive resources with which to engage in deep cognitive processing. The spatial contiguity effect (as well as several other effects explored in subsequent sections) can be seen as a specific example of the more general split attention effect. Importantly, patterns similar to our spatial contiguity effect have been reported by other researchers (Chandler & Sweller, 1991; Paas & Van Merrienboer, 1994; Sweller & Chandler, 1994; Sweller, Chandler, Tierney, & Cooper, 1990).

Further evidence comes from research by Hegarty, Carpenter, and Just (1996) concerning students' eye movements as they viewed a screen showing a diagram of a pulley system along with corresponding text (as in our integrated presentations). Students tended to read a sentence of text and inspect the corresponding portion of the illustration that was described in the text, then read the next sentence and inspect the corresponding portion of the illustration, and so on. This pattern of eye movement is consistent with the cognitive theory of multimedia learning in which students seek to represent and integrate corresponding small segments of text and illustrations. In our research, we attempt to foster this process by placing corresponding words and illustrations near each other on the page or screen.

V. Temporal Contiguity Effect

A. What Is the Temporal Contiguity Effect?

The foregoing section established a spatial contiguity effect in which students learn better when corresponding printed text and pictures are near rather than far from one another on the page or screen. In this section, we explore an analogous version of contiguity solely within a computer-based venue—namely, a temporal continuity effect in which students learn better when corresponding aspects of narration and animation are presented simultaneously rather than successively. Consider the narrated animation explaining the formation of lightning shown in Fig. 1. As you can see, the narration (indicated as a quotation under the frame) describes the same steps as are depicted in the concurrent animation (indicated in the graphic in the frame). For example, when the narrator says "negatively-charged particles fall to the bottom of the cloud" the animation simultaneously shows negative particles moving to the bottom of the cloud. This narrated animation has temporal contiguity because corresponding words and pictures are presented at the same time. We call this version simultaneous presentation. As an alternative that destroys temporal contiguity, we could present the entire narration followed by the entire animation (or vice versa). We call this version successive presentation. A temporal contiguity effect occurs if learners who receive simultaneous presentation perform better on transfer tests than do learners who receive successive presentation.

B. Theory: What Are the Mechanisms Underlying the Temporal Contiguity Effect?

The information delivery theory provides a case against temporal contiguity. According to a strict version of the information delivery theory, exactly the same information is delivered to the learner in simultaneous and successive presentations. Thus, the two groups should perform about the same on subsequent transfer tests. According to a lenient version of the information delivery theory, students in the successive group get the same message presented twice, whereas students in the simultaneous group get the message presented only once (in two different formats)—so the successive group gets twice as much exposure time as does the simultaneous group. Thus, the successive group should perform better on transfer tests than the simultaneous group.

The case for temporal contiguity comes from the cognitive theory of multimedia learning. According to the cognitive theory of multimedia learning, meaningful learning is most likely to occur when corresponding aspects of the verbal and pictorial representations are in working memory at the same time. Simultaneous presentation is more likely to facilitate this situation than is successive presentation, so the simultaneous group should perform better on transfer tests than the successive group.

C. Research: Is There a Temporal Contiguity Effect?

1. Core Findings

The top portion of Table XIII summarizes the effect sizes in comparing the advantage of simultaneous over successive presentation across eight studies involving

TABLE XIII

SUMMARY OF TEMPORAL CONTIGUITY EFFECTS: BETTER TRANSFER WHEN CORRESPONDING NARRATION AND ANIMATION ARE PRESENTED SIMULTANEOUSLY RATHER THAN SUCCESSIVELY [FOR ENTIRE PRESENTATION (A) BUT NOT FOR ALTERNATING SEGMENTS (B)]

Source	Medium	Content	Effect size
(A) Entire narration before or after entire animation			
Mayer & Anderson, 1991, Exp. 1	Screen	Pumps	1.00
Mayer & Anderson, 1991, Exp. 2	Screen	Pumps	1.05
Mayer & Anderson, 1992, Exp. 1	Screen	Pumps	1.61
Mayer & Anderson, 1992, Exp. 2	Screen	Brakes	1.33
Mayer & Sims, 1994, Exp. 1	Screen	Brakes	0.83
Mayer & Sims, 1994, Exp. 2	Screen	Lungs	1.60
Mayer et al., 1999, Exp. 1	Screen	Lightning	1.96
Mayer et al., 1999, Exp. 2	Screen	Brakes	1.27
Median			1.30
(B) Alternating short segments of corresponding narration and animation			
Mayer et al., 1999, Exp. 1	Screen	Lightning	.24
Mayer et al., 1999, Exp. 2	Screen	Brakes	.05
Moreno & Mayer, 1999, Exp. 2	Screen	Lightning	.12
Moreno & Mayer, 2001, Exp 2	Screen	Lightning	−.16
Median			.08

multimedia explanations of pumps (Mayer & Anderson, 1991, Experiments 1 and 2; Mayer & Anderson, 1992, Experiment 1), brakes (Mayer & Anderson, 1992, Experiment 2; Mayer & Sims, 1994, Experiment 1; Mayer, Moreno, Boire, & Vagge, 1999, Experiment 2), lungs (Mayer & Sims, 1994, Experiment 2), and lightning (Mayer et al., 1999, Experiment 1). As you can see, in eight out of eight comparisons, students who received narration and animation coordinated in time (i.e., simultaneous presentations) demonstrated better understanding of the explanation than did students who received the same narration and animation separated in time (i.e., successive presentation), yielding a median effect size of 1.30. This pattern is consistent with the cognitive theory of multimedia learning and inconsistent with the information delivery theory. We call these results the *temporal contiguity effect:* Students learn more deeply when corresponding segments of animation and narration are presented simultaneously rather than successively.

In a supplementary test of the cognitive theory of multimedia learning, we compared simultaneous presentation to successive presentation of short segments, that is, one sentence of script followed (or preceded) by a corresponding 10-s animation segment. According to the cognitive theory of multimedia learning, learners who receive successive presentation of short segments should perform as well as learners in the simultaneous presentation because they are able to hold corresponding segments of narration and animation in working memory at the same time. The results of four comparisons involving multimedia messages about lightning (Mayer et al., 1999, Experiment 1; Moreno & Mayer, 1999b, Experiment 2; Moreno & Mayer, in press, Experiment 2) or brakes (Mayer et al., 1999, Experiment 2) revealed that in four out of four comparisons there was no large difference between the groups, yielding a median effect size of .08. These results are summarized in the bottom of Table XIII and provide additional support for the cognitive theory of multimedia learning.

2. *Related Research*

A temporal contiguity effect was first reported by Baggett and her colleagues (Baggett, 1984, 1989; Baggett & Ehrenfeucht, 1983) in which students viewed a narrated film on how to use an assembly kit called Fischer Technik 50. The step-by-step procedure was shown on film and described in a voice overlay; the voice overlay was presented simultaneously with the corresponding portion of the film or preceding (or following) the corresponding film portion by 21 s. A temporal contiguity effect was found in which students who received simultaneous presentation performed better on subsequent assembly tasks than did students who received the same animation and narrative misaligned by 21 s.

VI. Coherence Effect

A. What Is the Coherence Effect?

The foregoing sections show that students have an easier time in building a coherent mental representation when corresponding words and pictures are presented together in time or space. In this section, I explore the idea that the knowledge construction process is facilitated when extraneous information is excluded from the presentation. For example, let us begin with a narrated animation (based on temporal contiguity) or a set of captioned illustrations (based on spatial contiguity). One way to modify these multimedia messages is to include or exclude additional sentences containing supporting information. For example, Fig. 4 shows illustrations with concise captions for the lightning lesson (which we call the concise presentation), and Table XIV lists longer captions for each illustration (which we call the expanded presentation). The expanded text is intended to support and clarify the core material found in the concise text.

TABLE XIV

EXPANDED TEXT FOR THE LIGHTNING LESSON

1. Warm moist air near the earth's surface rises. As the air in this updraft cools, water in this updraft cools, water vapor condenses into water droplets and forms a cloud. The cloud's top extends beyond the freezing level, so tiny ice crystals form in the upper portion of the cloud.
2. Eventually, the ice crystals become too large to be suspended by updrafts, so they fall through the cloud. They drag air downward, producing downdrafts. When downdrafts strike the ground they produce gusts of cool wind.
3. The rising water droplets collide with the falling ice, producing electrical charges. Negatively charged particles fall to the bottom of the cloud, and positively charged particles rise to the top.
4. A negatively charged stepped leader moves downward from the cloud in a series of steps. A positively charged upward-moving leader travels up from trees and buildings to meet the negative charges. When the two leaders meet, negatively charged particles rush from the cloud to the ground.
5. As the leader stroke nears the ground, it induces an opposite charge. Positively charged particles from the ground rush upward along the same path. This return stroke produces the bright light that people notice as a flash of lightning.

From Mayer et al. (1996).

Another way to modify these multimedia passages is to include or exclude interesting photos and facts about them with captioned illustrations (or interesting video clips and facts about them with narrated animation). For example, Table XV summarizes some video clips—each about 10 s in length—and accompanying narration segments—each about one sentence long—that were interspersed at appropriate places within the narrated animation about lightning formation (as shown in Fig. 1). The multimedia lesson with five interspersed video clips can be called the expanded presentation, whereas the one without added video clips can be called the concise presentation. The added video and narration does not replace the core material, but is intended to make the lesson more interesting.

Finally, a third way to modify multimedia messages is to include or exclude background sounds and music. For example, the narrated animation about lightning formation can be modified by adding environmental sounds such as blowing winds when "gusts of cool wind" are mentioned or cracking ice when "tiny ice crystals" are mentioned, and by adding a soft instrumental musical loop in the background. We refer to the version with added sounds as the expanded presentation and the version without added sounds as the concise presentation. The added sounds do not interfere with the narration but are intended to make the lesson more enjoyable.

In each of the three situations, if the concise group performs better than the expanded group, this would be an example of the *coherence effect*.

TABLE XV

INTERESTING VIDEO CLIPS AND CORRESPONDING NARRATION FOR LIGHTNING LESSON

Narration script	Video images
"Lightning can occur virtually any season and can potentially strike anywhere at any time."	Flashes of lightning in the sky and clouds, and above a cluster of trees.
"Scientists can simulate this process in a controlled laboratory experiment."	Swirling wind within glass-enclosed chamber.
"In trying to understand the various processes involved, scientists create lightning by launching rockets into overhead clouds."	Rockets being set up in open field, buttons being pressed on control box, small rockets soaring into clouds.
"Statistics show that more people are injured by lightning each year than by tornadoes and hurricanes combined."	Lighting strikes into trees, ambulance arrives, victim placed on stretcher, onlookers watch.
"When lightning strikes the ground, fulgurites may form as the heat from lightning fuses sand into the shape of the electricity's path."	Workers inspecting fulgurites, sweeping off sand with small brushes, applying metal instruments.
"Many people thought that lightning was a form of heavenly fire until Benjamin Franklin conducted his famous experiments with a kite and key showing that lightning was really a form of electricity."	Cloud-to-ground lightning strikes in various cities, from a panoramic skyline perspective.

B. THEORY: WHAT ARE THE MECHANISMS UNDERLYING THE COHERENCE EFFECT?

According to the information delivery theory, students who receive the expanded presentation and students who receive the concise presentation are exposed to identical information about how something works. Thus, both groups should perform about the same on subsequent transfer tests. The case for adding interesting and entertaining material comes from *arousal theory*—the idea that students are more likely to pay attention when they are emotionally aroused (Dewey, 1913; Renninger, Hidi, & Krapp, 1992). Adding interesting video clips or background sounds increases the learner's level of emotional arousal, which in turn, causes the learner to pay more attention to the incoming core information. Thus, the expanded group should outperform the concise group on subsequent tests of transfer.

In contrast, the cognitive theory of multimedia learning posits that adding interesting but irrelevant material can interfere with the learner's process of structure building by distraction (i.e., interfering with the *selecting* process by taking the reader's limited attentional resources away from the core material), disruption (i.e., interfering with the organizing process by putting extraneous material between the steps in the causal chain), or diversion (i.e., interfering with the *integrating* process by priming prior knowledge related to the added material). Harp and Mayer (1998) have provided evidence for the diversion hypothesis by showing that adding interesting but irrelevant pictures and text to a scientific explanation about lighting formation can encourage learners to integrate the incoming information with prior knowledge about the dangers of lightning rather than a cause-and-effect chain. Thus, the cognitive theory of multimedia learning predicts that the concise group should outperform the expanded group on transfer tests.

C. Research: Is There a Coherence Effect?

1. *Core Findings*

Does eliminating unneeded material improve learner understanding of a multimedia message? As shown in the three sections of Table XVI, we have examined this question in three ways—by seeing what happens when we add or eliminate irrelevant words from a book-based lesson on lightning (Mayer et al., 1996, Experiments 1, 2, and 3); irrelevant stories and pictures from a book-based lesson on

TABLE XVI

Summary of Coherence Effects: Better Transfer When Irrelevant Words, Pictures, and Sounds Are Excluded Rather Than Included

Source	Medium	Content	Effect size
Eliminate irrelevant words			
Mayer et al., 1996, Exp. 1	Page	Lightning	1.47
Mayer et al., 1996, Exp. 2	Page	Lightning	0.93
Mayer, et al., 1996, Exp. 3	Page	Lightning	2.31
Median			1.47
Don't add irrelevant words and pictures			
Harp & Mayer, 1997, Exp. 1	Page	Lightning	1.67
Harp & Mayer, 1998, Exp. 1	Page	Lightning	2.59
Harp & Mayer, 1998, Exp. 2	Page	Lightning	1.65
Harp & Mayer, 1998, Exp. 3	Page	Lightning	1.17
Harp & Mayer, 1998, Exp. 4	Page	Lightning	1.85
Mayer, Heiser, & Lonn, 2001, Exp. 1	Screen	Lightning	0.55
Median			1.66
Don't add irrelevant sounds and music			
Moreno & Mayer, 2000, Exp. 1	Screen	Lightning	1.27
Moreno & Mayer, 2000, Exp. 2	Screen	Brakes	0.96
Median			1.11

lightning (Harp & Mayer, 1997, Experiment 1; Harp & Mayer, 1998, Experiments 1, 2, 3, and 4) and a computer-based lesson on lightning (Mayer, Heiser, & Lonn, 2001, Experiment 1); and irrelevant sounds and music in a computer-based lesson on lightning (Moreno & Mayer, 2000a, Experiment 1) and brakes (Moreno & Mayer, 2000a, Experiment 2).

The top portion of Table XVI shows that in three out of three comparisons involving text with illustrations, eliminating irrelevant words from the text resulted in improved transfer test performance. The median effect size was 1.47. The middle portion of Table XVI shows that in six out of six comparisons involving text and illustrations or narration and animation, eliminating irrelevant facts from the text (or narration), and eliminating irrelevant photos (or video clips) from the graphics resulted in improved transfer test performance. The median effect size was 1.66. The bottom portion of Table XVI shows that in two out of two comparisons involving animation and narration, eliminating background sounds and music resulted in improved transfer test performance. The median effect size was 1.11. Overall, across 11 comparisons, students in the concise group performed better on transfer tests than did students in the expanded group.

This pattern of results is consistent with the cognitive theory of multimedia learning and inconsistent with information delivery theory. The results summarized in Table XVI provide clear and consistent support for the *coherence effect:* Students learn more deeply when extraneous words, pictures, and sounds are eliminated rather than included.

2. *Related Research*

Consistent with our first test of the coherence effect (i.e., comparing summaries to full presentations), Reder and Anderson (1980) found that college students remember more of the important material in a lesson after reading a summary than after reading an entire chapter. However, our research focuses on problem-solving transfer rather than retention as the major dependent measure. Consistent with our second test of the coherence effect (i.e., adding interesting but irrelevant pictures and words), researchers have found that adding *seductive details*—that is, interesting but irrelevant facts or stories—either hurts or does not help students' ability to remember the main information in a text passage (Garner, Gillingham, & White, 1989; Garner, Brown, Sanders, & Menke, 1992; Hidi & Baird, 1988; Mohr, Glover, & Ronning, 1984; Shirey, 1992; Shirey & Reynolds, 1988; Wade, 1992; Wade & Adams, 1990). Again, our research focuses on problem-solving transfer rather than retention as the best way to measure learner understanding. Finally, consistent with our third test of the coherence effect (i.e., adding background sounds and music), research on TV viewing shows that sound effects generally cause children to momentarily look at the screen rather than to reflect deeply on the presented material (Kozma, 1991).

VII. Modality Effect

A. What Is the Modality Effect?

Figure 9 shows selected frames and corresponding on-screen text from a multimedia presentation on lightning formation—which we call animation-and-text (AT) presentation. In contrast, an alternative presentation format is to present the same animation along with corresponding narration containing the same words as in Fig. 1 spoken at the same time—which we call animation-and-narration (AN)

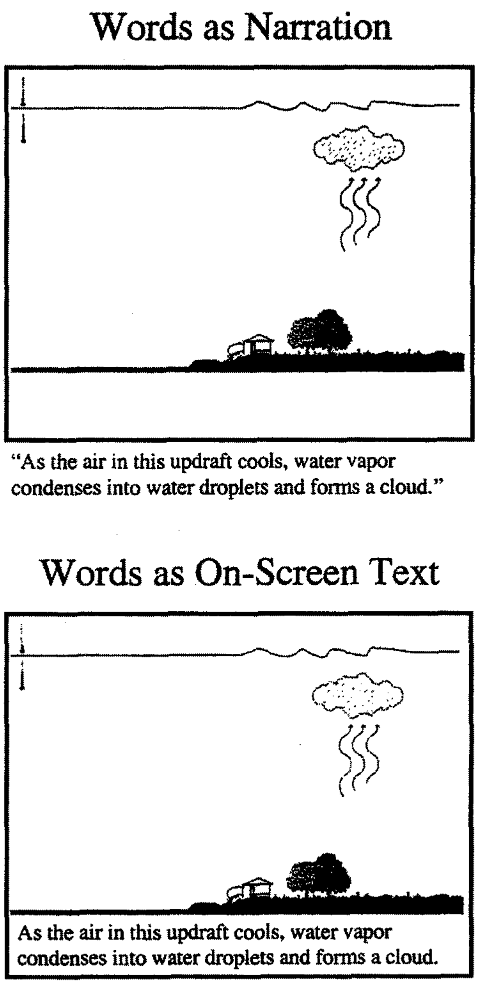


Fig. 9. Selected frames from a lightning lesson using animation and narration or animation and on-screen text.

presentation. In this section, I consider the cognitive consequences of presenting animation with corresponding narration (AN presentation) versus animation with corresponding on-screen text (AT presentation). In both cases, the animations are identical and the words are identical; and in both cases, the words are presented at the same time as the corresponding action in the animation. The only difference is whether the words are printed as on-screen text (AT presentation) or spoken as narration (AN presentation). Later, the learner takes a problem-solving transfer test involving open-ended questions, such as, "What could be done to reduce the intensity of a lightning storm?" or "Suppose you see clouds in the sky but not no lightning. Why not?" A *modality effect* (for transfer) occurs if students perform better on subsequent transfer tests when the words are spoken rather than printed, that is, when they receive animation and narration rather than animation and on-screen text.

In a computer game called Design-a-Plant (Lester, Towns, Callaway, Voerman, & Fitzgerald, 2000; Moreno, Mayer, & Lester, 2000), learners sitting at a computer screen or wearing a virtual reality head-mounted display take a simulated trip to a distant planet. An on-screen animated character (named Herman the Bug) describes the conditions on the planet (e.g., it's rainy and windy) and asks the learner to design a plant capable of surviving in that environment (e.g., by choosing appropriate roots, stems, and leaves). The on-screen character provides useful feedback that explains how plants grow, in a series of narrated animations. For some learners, Herman's words are presented as speech (including animation with text), whereas for other learners, Herman's words are presented as on-screen text (including animation with text). The learner takes trips to several planets and later is tested on problem-solving transfer problems, such as designing plants for new planets or telling for which environments a particular plant is best suited.

B. Theory: What Are the Mechanisms Underlying the Modality Effect?

According to an information delivery theory, the AN and AT presentations should result in identical learning outcomes because identical information is presented—namely, the same pictures and the same words. On the contrary, based on the cognitive theory of multimedia learning, we can predict that the AN presentation will lead to deeper learning than the AT presentation. The top row of Fig. 10 shows the sequence of cognitive processing for the AT presentation. The animation and text both enter the learner's cognitive system through the eyes and must be processed, at least initially, within visual working memory. Therefore, the AT presentation is more likely to cause overload in the visual channel because both words and pictures must compete for limited resources in visual working memory. The bottom row of Fig. 10 shows the sequence of cognitive processing for the AN presentation. The animation enters the learner's cognitive system through the eyes and is processed, at least initially, in visual working memory; the narration

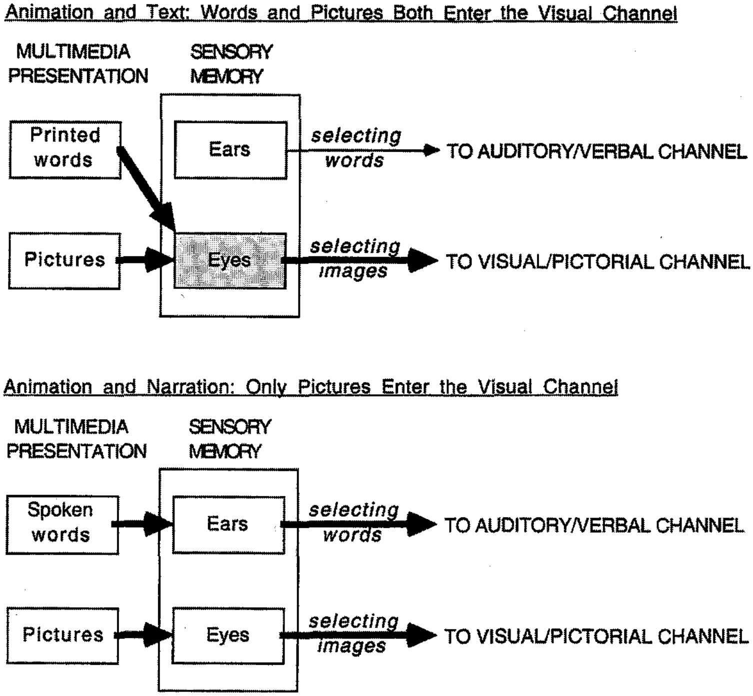


Fig. 10. Cognitive processing for animation with narration and for animation with on-screen text. (Adapted from Mayer 2001. Reprinted by permission. Copyright 2001 Cambridge University Press.)

enters the learner's cognitive system through the ears and is processed, at least initially, in auditory working memory. In contrast to the AT presentation in which all information is processed in the visual channel, in the AN presentation the verbal information can be off-loaded from the visual channel to the verbal channel. By off-loading the verbal information to the verbal channel, there are more cognitive resources available for processing the animation in visual working memory, thus reducing cognitive load. The cognitive consequences of the AN presentation include reduced cognitive load for representing the incoming information, thus allowing more cognitive capacity for making connections—the key component in deep learning.

C. Research: Is There a Modality Effect?

1. *Core Findings*

Does modality affect learning? To help answer this question, I identified 12 experiments in which we compared the transfer performance of students who had received an AN presentation and those who had received an AT presentation within

TABLE XVII

SUMMARY OF MODALITY EFFECTS: BETTER TRANSFER FROM ANIMATION AND NARRATION THAN FROM ANIMATION AND ON-SCREEN TEXT

Source	Medium	Content	Effect size
Moreno & Mayer, 1998, Exp. 1	Screen	Lightning	1.68
Moreno & Mayer, 1998, Exp. 2	Screen	Brakes	0.94
Moreno & Mayer, 1999, Exp. 1	Screen	Lightning	1.06
Moreno & Mayer, 1999, Exp. 2	Screen	Lightning	1.28
O'Neil et al., 2000, Exp. 1	Game	Aircraft	0.97
Moreno & Mayer, 2001, Exp. 1	Game	Plants	0.93
Moreno & Mayer, 2001, Exp. 2	Game	Plants	0.70
Moreno & Mayer, 2001, Exp. 3	Game	Plants	2.08
Moreno et al., in press, Exp. 4a	Game	Plants	0.51
Moreno et al., in press, Exp. 4b	Game	Plants	1.43
Moreno et al., in press, Exp. 5a	Game	Plants	1.20
Moreno et al., in press, Exp. 5b	Game	Plants	1.56
Median			1.13

the context of a short explanation of the process of lightning formation (Mayer & Moreno, 1998, Experiment 1; Moreno & Mayer, 1999b, Experiments 1 and 2), a short explanation of how a car's braking system works (Mayer & Moreno, 1998, Experiment 2), a virtual reality game explaining how an aircraft's fuel system works (O'Neil et al., 2000, Experiment 1), a computer game explaining plant growth (Moreno & Mayer, 2001, Experiment 1; Moreno et al., in press, Experiments 4a, 4b, 5a, and 5b), and a virtual reality game explaining plant growth (Moreno & Mayer, 2001, Experiments 2 and 3). Table XVII shows the effect sizes on transfer tests comparing the AN and AT groups across the 12 studies. In 12 out of 12 comparisons, the AN group performed better on solving transfer problems than did the AT group, yielding a median effect size of 1.13. Even though all students received the same animation along with the same words, students understood the material more deeply when the words were spoken rather than printed.

These results provide clear and consistent evidence for a *modality effect:* Students learn more deeply from animation and narration than from animation and text. The modality effect is consistent with the predictions of the cognitive theory of multimedia learning and inconsistent with the predictions of the information delivery theory.

2. *Related Research*

Although the studies listed in Table XVII were the first to report modality effects involving computer-based presentations, the first modality effect involving paper-based presentations was reported by Mousavi, Low, and Sweller (1995). Students received worked-out examples for solving geometry problems, presented as sheets of paper containing printed diagrams showing a step-by-step solution. In addition to the diagrams, some students also listened to a tape recording explaining how to solve the problem (narration and illustration), whereas other students received the same words as printed text (printed text and illustration). Students who received narration and illustration performed better on subsequent tests of geometry problem solving than did students who received printed text and illustration. This modality effect was replicated across several experiments (Mousavi et al., 1995), and similar results were obtained with other materials (Tindall-Ford, Chandler, & Sweller, 1997). In a review of research on the modality effect, Sweller (1999) proposed that audiovisual presentations are likely to be ineffective when the text is complex or when the visual material can be easily understood without the text.

VIII. Redundancy Effect

A. What Is the Redundancy Effect?

So far, our research shows that students can learn deeply from a concise narrated animation (such as in Fig. 1 for the lightning lesson). However, in an attempt to improve on a concise narrated animation, a designer might be tempted to add on-screen text, thus allowing learners the option of reading or listening to the words. Figure 11 shows selected frames from the lightning lesson consisting of animation (indicated by the graphics in each frame), narration (indicated by the quotation under each frame), and on-screen text (indicated by the text at the bottom of each frame). If students in the animation-with-narration-and-text group perform better on transfer tests than students in the animation-with-narration group, we have found a way to improve on concise narrated animation. If students in the animation-with-narration-and-text group perform worse on transfer tests than students in the animation-with-narration group, we have found a *redundancy effect,* in which adding redundant text to a narrated animation hurts learning.

B. Theory: What Are the Mechanisms Underlying the Redundancy Effect?

The case for adding on-screen text comes from a version of the information delivery theory. According to the information delivery theory, adding on-screen text to a narrated animation should either help or not hurt learning. Each mode of presentation—animation, narration, and on-screen text—is a vehicle for delivering the information in the lesson. In the strictest sense, adding extra delivery vehicles

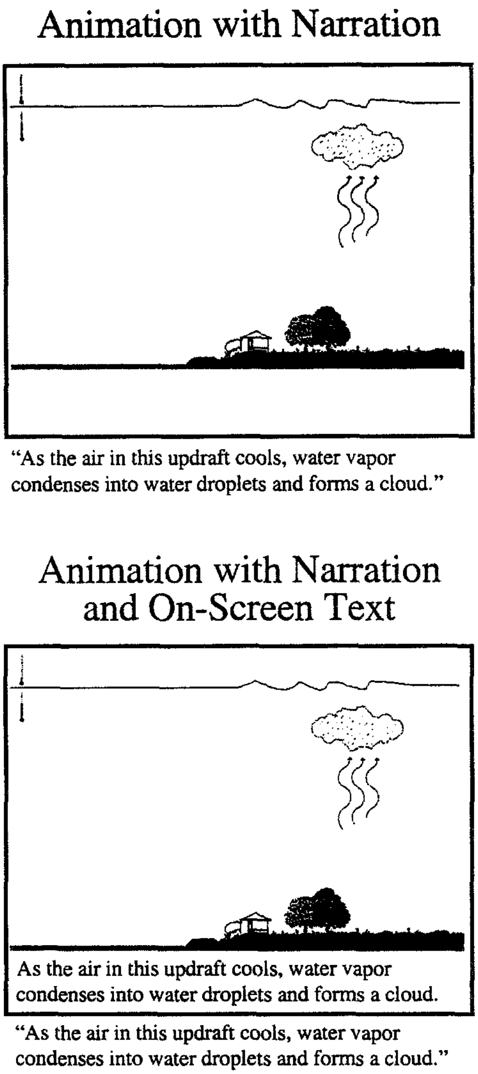


Fig. 11. Selected frames from a lightning lesson using animation, narration, and on-screen text.

should have no effect because any one can do the job. Thus the animation-with-narration-and-text group should perform like the animation-with-narration group on transfer tests.

In a more lenient interpretation—which can be called the learning preferences hypothesis (Plass, Chun, Mayer, & Leutner, 2000)—adding extra delivery vehicles should improve learning, especially if some paths to the learner are blocked. Adding extra delivery modes enables the learner to focus on the mode that corresponds to the learner's preferences. For example, if some learners prefer to learn from spoken words and others prefer to learn from printed words, then animation-with-narration-and-text can accommodate both kinds of learners. In contrast, animation-with-narration would not accommodate learners who prefer to learn from printed words. Thus, the animation-with-narration-and-text group should perform better than the animation-with-narration group on transfer tests.

The case against adding on-screen text is supported by the cognitive theory of multimedia learning. According to the cognitive theory of multimedia learning, adding on-screen text to a narrated animation can overload the visual channel because on-screen text and animation must both be processed through the eyes. In contrast, when only animation and narration are presented, the animation enters through the eyes and the narration enters through the ears, thus minimizing the cognitive load in each channel (Kalyuga, Chandler, & Sweller, 1998; Sweller, 1999). The remaining cognitive resources can be used for mentally building connections among the representations, an activity that leads to deeper learning. Thus, the animation-with-narration-and-text group should perform worse on transfer tests than the animation-with-narration group.

C. Research: Is There a Redundancy Effect?

1. Core Findings

Does redundancy affect learning? Three of our research studies provide relevant evidence in which we compare the transfer test performance of students who learned about lightning formation with animation and narration versus those who learned with animation, narration, and on-screen text (Mayer, Heiser, & Lonn, 2001, Experiments 1 and 2; Moreno & Mayer, in press, Experiment 2). In all three comparisons, summarized in Table XVIII, students in the animation-with-narration group outperformed those in the animation-with-narration-and-text group, yielding

TABLE XVIII

SUMMARY OF REDUNDANCY EFFECTS: BETTER TRANSFER FROM ANIMATION AND NARRATION THAN FROM ANIMATION, NARRATION, AND ON-SCREEN TEXT

Source	Medium	Content	Effect size
Mayer, Heiser, & Lonn, 2001, Exp. 1	Screen	Lightning	0.84
Mayer, Heiser, & Lonn, 2001, Exp. 2	Screen	Lightning	1.65
Moreno & Mayer, 2001, Exp. 2	Screen	Lightning	0.69
Median			0.84

a median effect size of .84. In these studies, adding more delivery modes resulted in poorer learning.

These results provide clear and consistent evidence for a *redundancy effect:* Students learn more deeply from animation and narration than from animation, narration, and text. The redundancy effect is consistent with the cognitive theory of multimedia learning and inconsistent with both versions of the information delivery theory.

2. *Related Research*

Complementary results were obtained in a study by Kalyuga, Chandler, and Sweller (1999) in which trainees learned about soldering metals in an industrial setting. Some people learned from diagrams accompanied by spoken instructions and some people learned from diagrams accompanied by spoken instructions and printed text (which contained the same words as the speech). There was a redundancy effect in which people who learned from diagrams and speech performed better on subsequent soldering tests than did those who learned from diagrams, speech, and printed text.

Kalyuga et al. (1998, p. 2) define the redundancy effect broadly to refer to any situation in which "eliminating redundant materials results in better performance than when the redundant material is included," and they provide a review of supporting research evidence. In this chapter, I use the term redundancy effect more narrowly to refer to situations in which adding on-screen text to a narrated animation results in poorer learning.

IX. Pretraining Effect

A. What is the Pretraining Effect?

Let us begin with a concise narrated animation such as the lesson on lightning (in Fig. 1), on brakes (in Fig. 2), or on pumps (in Fig. 3). What can we do to help learners build a mental model of how the system works, that is, a cause-and-effect model in which a change in one part causes a change in the next part and so on? Inexperienced learners may lack appropriate knowledge about the components in the system. For example, they may not know what a piston is or how it moves (in the brakes lesson) or what an inlet valve is and how it moves (in the pumps lesson).

To help learners make sense of the concise narrated animation, it might be useful for them to have some previous experience with the major components that are mentioned in the narration. For example, in the narrated animation for brakes (shown in Fig. 2) learners need to recognize the brake pedal, piston in master cylinder, fluid in tubes, smaller pistons in wheel cylinder, brake shoe, and brake

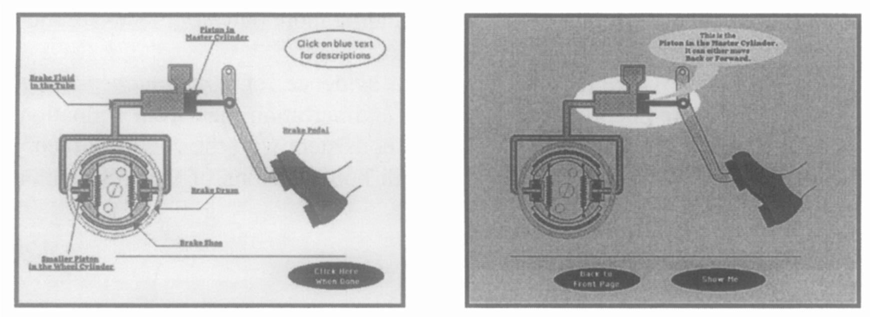


Fig. 12. Selected frames from a self-paced pretraining lesson on the components of a braking system.

drum. Figure 12 shows frames from a self-paced pretraining lesson intended to allow the learner to recognize and name each main part and know its behavior (i.e., the states that the part can be in). On the screen the learner sees a diagram of the braking system with a blue label next to each major part. When the learner clicks on the label for a part, the part is highlighted in an oval frame and a caption appears describing the possible states of the part (e.g., "This is the piston in the master cylinder. It can either move forward or back."). The learner can click on a "SHOW ME" button to see how the part moves (with all other parts of the system blacked out). Once the learner has explored each of the six parts of the braking system, the learner is ready for the narrated animation describing how the braking system works. We refer to this sequence of self-paced parts training followed by training on the whole system as pretrained presentation; when the self-paced parts training follows the narrated animation we call this posttrained presentation. A pretraining effect occurs when learners in the pretrained group outperform learners in the posttrained group on tests of transfer.

B. Theory: What Are the Mechanisms Underlying the Pretraining Effect?

According to the information delivery theory, learners receive exactly the same information in the pretrained and posttrained groups, so they should perform the same on subsequent tests of learning.

In contrast, the cognitive theory of multimedia learning views meaningful learning as a process of mental model construction. The process of mental model construction may create a heavy cognitive load during learning so inexperienced learners may have difficulty (Chi, 2000; Gentner & Stevens, 1983). Following earlier work (Bobrow, 1985; Gentner & Stevens, 1983; Mayer & Chandler, 2001), we propose a two-stage theory of mental model construction in which the learner builds *component models*—an understanding of the behavior, location, and name of each part—followed by a *causal model*—an understanding of the cause-and-effect chain of how a change in one part affects a change in another part and so on. When the learner receives a narrated animation, the learner must build component models and a causal model at the same time—a task that can overload the cognitive system. For example, when the lesson describes a piston moving forward in the master cylinder, the learner needs to be able to look at the portion of the animation depicting the piston in the master cylinder and recognize that it is moving forward. A way to reduce the load is to provide previous experience with the components so that the learner does not have to build component models while processing the narrated animation. In this way the learner can devote his or her full attention to building a causal model, resulting in deeper learning. Accordingly, we can predict that pretrained learners will learn more deeply from a narrated animation than those who have not received pretraining.

C. Research: Is There a Pretraining Effect?

1. *Core Findings*

Table XIX summarizes three comparisons of pretrained and posttrained groups. In the first study (Mayer & Chandler, 2001, Experiments 1), learners received a concise narrated animation about lightning formation either before or after viewing the same presentation segment-by-segment by clicking on a button to go on to the next segment. In the second study (Mayer & Mathias, 2001, Experiment 2), learners received a concise narrated animation about brakes either before or after interacting with a multimedia description of each part of in the braking system controlled by the learner clicking on a diagram showing each part. In the third study (Mayer & Mathias, 2001, Experiment 3), learners received a concise narrated animation about a bicycle tire pump either before or after interacting with a concrete model of the pump demonstrating the operation of each part. As you can see in the table,

TABLE XIX

Summary of Pretraining Effects: Better Transfer When Training on Components Precedes Rather than Follows a Message

Source	Medium	Content	Effect size
Mayer & Chandler, 2001, Exp. 1	Screen	Lightning	1.14
Mayer & Mathias, 2001, Exp. 2	Screen	Brakes	1.39
Mayer & Mathias, 2001, Exp. 3	Screen	Pumps	2.16
Median			1.39

in three out of three comparisons, learners who received pretraining about each part of the system before they received the concise narrated animation performed better on transfer tests than did learners who received the same training about parts after the concise narrated animation. The median effect size was 1.39. These results provide support for the cognitive theory of multimedia learning and against the information delivery theory. In sum, across three comparisons, we found evidence for the *pretraining effect:* Students learn more deeply when training on components precedes rather than follows the multimedia message.

2. *Related Research*

Previous research on learner control and pacing of computer-based instruction has yielded mixed results (Lepper, 1985; Williams, 1996). We overcame this problem by focusing on a very specific aspect of learner control in multimedia learning—by examining the placement of self-paced training on the components in a to-be-learned system (i.e., before or after the multimedia lesson).

X. Signaling Effect

A. What Is the Signaling Effect?

Let us begin with a narrated animation that explains how airplanes achieve lift, based on the ideas that the upper end of the wing is more curved (and therefore has more surface area) than the bottom of the wing, that air moves faster over the top than the bottom of the wing, and pressure is less on the top than the bottom of the wing. The narrated animation contains about 600 words and runs for about 4 min (Mautone & Mayer, 2001). We call this a nonsignaled presentation because it lacks an introductory outline, headings, and pointer words (such as "first … second … third") that highlight the structure of ideas.

In an attempt to guide the learner's processing of the material, we can add signaling to the narration (Lorch, 1989; Meyer, 1975)—consisting of a 69-word preview paragraph that outlines the three main ideas in the passage (about wing shape, air speed, and air pressure), three headings to mark the sections on wing shape, air speed, and air pressure), and 16 connecting words such as "as a result" and "because it's curved." The signals do not add any additional content but rather are intended to highlight how the ideas are organized into a causal chain. We refer to this version in which the narration is signaled as a signaled presentation.

We can test students by asking them to write answers to transfer questions such as, "How can a plane be designed to achieve lift more rapidly?" or "Using what you've learned about how airplanes achieve lift, explain how helicopters achieve lift." A signaling effect would be indicated if students perform better on generating appropriate answers on the transfer questions after having received a signaled rather than nonsignaled presentation.

B. Theory: What Are the Mechanisms Underlying the Signaling Effect?

According to the information delivery theory, both the signaled and nonsignaled groups receive exactly the same information, so both groups should perform about the same on subsequent tests of learning.

In contrast, the cognitive theory of multimedia learning posits that signaling can guide the learner's cognitive processing during learning—particularly, the selecting of relevant words and the organizing of the words into a coherent cause-and-effect chain. Although no new content is present within the signals, the signals are intended to convey a sense of which ideas are important and how they are related to one another. For the airplane lift lesson, the signals point out the key steps in the causal chain and show how one is related to the next. Thus, the cognitive theory of multimedia learning predicts that narrated animations with signaled narration should lead to better transfer performance than narrated animations without signaled narration.

C. Research: Is There a Signaling Effect?

1. *Core Findings*

Table XX summarizes the results of two comparisons between multimedia messages in which the narration was signaled versus nonsignaled. In both messages, the narrated animation explained how airplanes achieve lift (Mautone & Mayer, 2001, Experiments 3a and 3b). As you can see, in both comparisons learners who received signaled narration performed better on transfer tests than did learners who received nonsignaled narration, with a median effect size of .60. This pattern of results provides a moderate level of initial support for the *signaling effect:* Students learn more deeply when the narration in a multimedia message is signaled rather than nonsignaled.

TABLE XX

Summary of Signaling Effects: Better Transfer When Narration is Signaled Rather than Nonsignaled

Source	Medium	Content	Effect size
Mautone & Mayer, 2001, Exp 3a	Screen	Airplane lift	0.58
Mautone & Mayer, 2001, Exp. 3b	Screen	Airplane lift	0.62
Median			0.60

2. Related Research

Research on signaling of printed text in a book-based environment shows that students who read signaled text performed better on retention or transfer tests than did students who received nonsignaled text (Loman & Mayer, 1983; Lorch, 1989; Rickards, Fajen, Sullivan, & Gillespie, 1997). More recently, Mautone and Mayer (2001) also found that students performed better on transfer tests when they read a signaled text rather than a nonsignaled text explaining how airplanes achieve lift (effect size = .74) or listened to a signaled narration rather than a nonsignaled narration explaining how airplanes achieve lift (effect size = .76). Overall, Mautone and Mayer's research extends the study of signaling to a multimedia learning environment.

XI. Personalization Effect

A. What Is the Personalization Effect?

The foregoing effects are all consistent with a cognitive theory of multimedia and provide strong support for the use of concise narrated animation. In our search for ways to improve on concise narrated animation, we examined the role of conversational style. In particular, suppose we change the narration in the lightning message from its formal textbook-like style to a more personalized conversational style, such as shown in Table XXI. As you can see, the personalized version contains the same factual information as the nonpersonalized version. However,

TABLE XXI

Portion of Personalized Narration for the Lightning Passage

Let me tell you what happens when lightning forms. Suppose you are standing outside, feeling the warm rays of sun heating up the earth's surface around you. Cool moist air moves over a warmer surface and becomes heated. The warmed moist air near the earth's surface rises rapidly. As the air in this updraft cools, water vapor condenses into water droplets and forms a cloud. Congratulations! You have just witnessed the birth of your own cloud.

As you watch, you tilt your head skyward. Your [The] cloud's top extends above the freezing level, so the upper portion of your [the] cloud is composed of tiny ice crystals. Brrr! I'm feeling cold just thinking it! Eventually, the water droplets and ice crystals become too large to be suspended by updrafts. As raindrops and ice crystals fall through your [the] cloud, they drag some of the air in your [the] cloud downward, producing downdrafts. When downdrafts strike the ground, they spread out in all directions, producing the gusts of cool wind you [people] feel just before the start of the rain. If you could look inside your cloud, you could see a neat pattern: Within the cloud, the rising and falling air currents cause electrical charges to build. The negatively charged particles fall to the bottom of the cloud, and most of the positively charged particles rise to the top.

Note. Underlined portions were added to the personalized version (or replaced the bracketed word).

unlike the formal third-person style of the nonpersonalized version, the personalized version contains phrases presented in the first and second person as if the narrator were talking directly to the learner in a conversation. A *personalization effect* occurs if people learn more deeply when narrators speak (or authors write) in personalized conversational style rather than nonpersonalized formal style.

B. Theory: What Are the Mechanisms Underlying the Personalization Effect?

The information delivery theory provides the rationale for the case against personalized instructional messages. If both personalized and nonpersonalized messages contain the same information, then transfer test performance should be about the same with both types of presentations.

In contrast, we can amend the cognitive theory of multimedia learning to include social factors that influence the learner's effort in engaging in deep cognitive processing (such as constructing and integrating visual and verbal models). Building on Reeves and Naas' (1996) *media equation hypothesis*—the idea that people easily accept computers as social partners—we view personalization as a technique that can encourage learners to react to the computer as a social agent. Long-standing theories of conversation (Grice, 1975) posit that people involved in conversation do so on the basis of certain conversational rules, including a commitment to try to understand what the other speaker is saying. If the speaking style of a computer-based narrator primes the conversational schema in a learner, then the learner is more likely to try hard to understand the speaker's explanation as would be done in human-to-human conversation. If personalized messages prime the conversation schema in learners, then transfer test performance should be better for personalized than for nonpersonalized messages. On the contrary, if personalization results in the addition of distracting irrelevant material, then transfer performance should be poorer for personalized than nonpersonalized messages (as with the coherence effect).

C. Research: Is There a Personalization Effect?

1. Core Findings

Do students learn more deeply from personalized rather than nonpersonalized multimedia messages? We have investigated this issue in five separate studies involving learning about the process of lightning formation in a computer environment (Moreno & Mayer, 2000, Experiments 1 and 2) and learning about the design of plants in a game environment (Moreno & Mayer, 2000, Experiments 3, 4, and 5). Table XXII summarizes the effect sizes attributable to personalization in each study. As you can see, there is a strong and consistent personalization effect, yielding a median of 1.55. On average, students who learn from an agent who speaks in personalized style perform 1.5 standard deviations better on a transfer

TABLE XXII

SUMMARY OF PERSONALIZATION EFFECTS: BETTER TRANSFER WHEN WORDS ARE IN CONVERSATIONAL STYLE RATHER THAN FORMAL STYLE

Source	Medium	Content	Effect size
Moreno & Mayer, 2000, Exp. 1	Screen	Lightning	0.96
Moreno & Mayer, 2000, Exp. 2	Screen	Lightning	1.60
Moreno & Mayer, 2000, Exp. 3	Game	Plants	1.55
Moreno & Mayer, 2000, Exp. 4	Game	Plants	1.59
Moreno & Mayer, 2000, Exp. 5	Game	Plants	0.88
Median			1.55

test than do students who learn from an agent who speaks in a nonpersonalized style. The *personalization effect* is that students learn more deeply when words are presented in conversational style rather than formal style. The results are consistent with a social agency adaptation of the cognitive theory of multimedia learning in which social cues in a multimedia message prime a social conversation schema in learners, leading to deeper cognitive processing.

2. *Related Findings*

In a related set of studies recently carried out in our lab we replicated the personalization effect using a different passage—a 45-s narrated animation explaining how the human respiratory system works. In particular, students who received a narrated animation spoken in conversational style generated more correct answers on a transfer test than did students who received the same material spoken in a formal style (effect size = 1.69). We have also recently found that students perform better on transfer tests when the voice is from a nonaccented male than when it is from a male who has a foreign accent (effect size = .80) or when it is from a machine-simulated voice (effect size = .81). These results show the power of voice and conversational style as social cues that prime social conversational schemas in learners.

XII. Other Effects

Our research also shows that improvements in the design of multimedia instructional messages have stronger effects for low-knowledge rather than high-knowledge learners (Mayer & Gallini, 1990; Mayer et al., 1995) and for high-spatial ability rather than low-spatial ability learners (Mayer & Sims, 1994).

We refer to these findings as *individual differences effects* (Mayer, 2001b) and note that additional research is needed to establish their robustness. High-knowledge learners may be able to compensate for poorly designed presentations by mentally rearranging them, whereas low-knowledge learners are less able to mentally repair poorly designed presentations. Low-spatial ability learners may have to allocate so much cognitive effort to building and holding mental images that they are unable to benefit from well-designed presentations; in contrast, high-spatial ability learners may have the cognitive capacity available to carry out deep cognitive processing fostered by well-designed presentations.

XIII. Conclusion

Table XXIII summarizes nine multimedia learning effects we have discovered in our research on multimedia learning. For each effect, we provide a short

TABLE XXIII

NINE MULTIMEDIA LEARNING EFFECTS

Principle	Number of tests	Median effect size
Multimedia effect: Better transfer when a message contains words and pictures rather than words alone.	11 of 11	1.39
Spatial contiguity effect: Better transfer when printed words are placed near rather than far from corresponding pictures.	5 of 5	1.12
Temporal contiguity effect: Better transfer when corresponding narration and animation are presented simultaneously rather than successively.	8 of 8	1.30
Coherence effect: Better transfer when irrelevant words, pictures, and sounds are excluded rather than included.	11 of 11	1.11
Modality effect: Better transfer from animation and narration than from animation and on-screen text.	12 of 12	1.13
Redundancy effect: Better transfer from animation and narration than from animation, narration, and on-screen text.	3 of 3	0.84
Pretraining effect: Better transfer when training on components precedes rather than follows a message.	3 of 3	1.39
Signaling effect: Better transfer when narration is signaled rather than nonsignaled.	2 of 2	0.60
Personalization principle: Better transfer when words are in conversational style rather than formal style.	5 of 5	1.55

description, tell how many comparisons we made, and report the median effect size. Concerning consistency of results, our research base allowed us to make 60 comparisons, with each comparison supporting the predictions of the cognitive theory of multimedia learning. Concerning the significance of results, most of the effects yielded median effect sizes greater than 1. Overall, our empirical results appear to be consistent and strong. This outcome satisfies our empirical goal of building a solid research for theory and practice.

Taken together, the nine effects allow us to test—and in some cases to clarify—a cognitive theory of multimedia learning. Each effect provides an independent source of evidence for the cognitive theory of multimedia learning and its core assumptions concerning dual channels, limited capacity, and active cognitive processing. This outcome satisfies our theoretical goal of contributing to a cognitive theory of how people learn from words and pictures.

Taken together, the nine effects also provide practical guidance for the design of multimedia instructional messages (Clark, 1999; van Merrienboer, 1997). Our principles (Mayer, 2001b) are most relevant when the designer's goal is to construct book-based, computer-based, or game-based explanations of how something works for naive learners. This outcome satisfies our practical goal of developing principles for multimedia design based on theory-based research rather than the intuitions of designers.

Our work is limited by the nature of the multimedia messages we employed (short causal explanations), the nature of our dependent measure (problem-solving transfer), the nature of our participants (high-school and college students who generally were unfamiliar with the material), and the learning context (a psychology laboratory). I focused on how naive learners come to understand scientific explanations because this is a central challenge of science education and because of my larger interest in how to promote problem-solving transfer (Mayer, 2002; Mayer & Wittrock, 1996). Further research is needed to determine the extent to which our laboratory-derived findings apply in more authentic classroom situations.

Finally, this report on our program of research can be seen as a case example of how practical educational problems can challenge cognitive psychologists to develop more authentic theories of how people learn (Bransford et al., 1999; Lambert & McCombs, 1998). Consistent with the evolution of learning theory in the 20th century (Mayer, 2001a), theories of learning in the 21st century can be strengthened when they are challenged to account for how people learn in practical educational situations—such as multimedia learning environments.

ACKNOWLEDGMENTS

I gratefully acknowledge the following colleagues who have contributed substantially to our research program on multimedia learning: Richard B. Anderson, Paul Chandler, Dorothy Chun, Joan

Gallini, Shannon Harp, Julie Heiser, Detlev Leutner, Steve Lonn, Patricia Mautone, Roxana Moreno, Harry O'Neil, Jan Plass, and Valerie Sims.

Portions of this chapter are adapted from Mayer, R. E. (2001). *Multimedia Learning*. New York: Cambridge University Press. The author's mailing address is: Richard E. Mayer, Department of Psychology, University of California, Santa Barbara, CA 93106. The author's email address is: mayer@psych.ucsb.edu.

REFERENCES

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* New York: Longman.

Ausubel, D. P. (1968). *Educational psychology: A cognitive view.* New York: Hott, Rinehart & Winston.

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

Baddeley, A. D. (1999). *Human memory.* Boston: Allyn & Bacon.

Baggett, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media associations. *Journal of Educational Psychology, 76,* 408–417.

Baggett, P. (1989). Understanding visual and verbal message. In H. Mandl & J. R. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 101–124). Amsterdam: Elsevier.

Baggett, P., & Ehrenfeucht, A. (1983). Encoding and retaining information in the visuals and verbals of an educational movie. *Educational Communications and Technology Journal, 31,* 23–32.

Bartlett, F. C. (1932). *Remembering.* London: Cambridge University Press.

Bobrow, D. G. (Ed.), (1985). *Qualitative reasoning about physical systems.* Cambridge, MA: MIT Press.

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

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

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology, 62,* 233–246.

Chi, M. T. H. (2000). Self-explaining: The dual processes of generating inference and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology: Volume 5, Educational design and cognitive science* (pp. 161–238). Mahwah, NJ: Erlbaum.

Clark, R. C. (1999). *Developing technical training* (2nd ed). Washington, DC: International Society for Performance Improvement.

Cognition and Technology Group at Vanderbilt (1996). Looking at technology in context: A framework for understanding technology and education. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 807–840). New York: Macmillan.

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920.* New York: Teachers College Press.

Dewey, J. (1913). *Interest and effort in education.* Cambridge, MA: Houghton Mifflin.

Ebbinghaus, H. (1995/1964). *Memory.* New York: Dover. [Originally published in 1885.]

Garner, R., Brown, R., Sanders, S., & Menke, D. (1992). Seductive details and learning from text. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 239–254). Hillsdale, NJ: Erlbaum.

Garner, R., Gillingham, M., & White, C. (1989). Effects of seductive details on macroprocessing and microprocessing in adults and children. *Cognition and Instruction, 6,* 41–57.

Gentner, D., & Stevens, A. L. (1983). *Mental models.* Mahwah, NJ: Erlbaum.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3, Speech acts* (pp. 51–58). New York: Seminar Press.

Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology, 89,* 92–102.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90,* 414–434.

Hegarty, M., Carpenter, P. A., & Just, M. A. (1996). Diagrams in the comprehension of scientific texts. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research.* Volume II (pp. 641–668). Mahwah, NJ: Erlbaum.

Hidi, S., & Baird, W. (1988). Strategies for increasing text-based interest and students' recall of expository text. *Reading Research Quarterly, 23,* 465–483.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40,* 1–17.

Kalyuga, S., Chandler, P., & Sweller, P. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13,* 351–372.

Kintsch, W. (1998). *Comprehension.* Cambridge, UK: Cambridge University Press.

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

Lachman, R., Lachman, J. T., & Butterfield, E. C. (1979). *Cognitive psychology and information processing.* Mahwah, NJ: Erlbaum.

Lambert, N. M., & McCombs, B. L. (1998). *How students learn.* Washington, DC: American Psychological Association.

Lepper, M. R. (1985). Microcomputers in education: Motivational and social issues. *American Psychologist, 40,* 1–18.

Lester, J. C., Towns, S. G., Callaway, C. B., Voerman, J. L., & Fitzpatrick, P. J. (2000). Deictic and emotive communication in animated pedagogical agents. In J. Cassell, J. Sullivan, S. Prevost, & E. Churchill (Eds.), *Embodied conversational agents* (pp. 123–154). Cambridge, MA: MIT Press.

Loman, N. L., & Mayer, R. E. (1983). Signaling techniques that increase the understandability of expository prose. *Journal of Educational Psychology, 75,* 402–412.

Lorch, R. F. (1989). Text signaling devices and their effects on reading and memory processes. *Educational Psychology Review, 1,* 209–234.

Mandl, H., & Levin, J. R. (Eds.). (1989). *Knowledge acquisition from text and pictures.* Amsterdam: North-Holland.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a guide in multimedia learning. *Journal of Educational Psychology, 93,* 377–389.

Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology, 81,* 240–246.

Mayer, R. E. (1992). Cognition and instruction: Their historic meeting within educational psychology. *Journal of Educational Psychology, 84,* 405–412.

Mayer, R. E. (1993). Illustrations that instruct. In R. Glaser (Ed.), Advances in instructional psychology (Vol. 4, pp. 253–284). Hillsdale, NJ: Erlbaum.

Mayer, R. E. (1996). Learners as information processors: Legacies and limitations of educational psychology's second metaphor. *Educational Psychologist, 31,* 151–161.

Mayer, R. E. (1999a). Instructional technology. In F. T. Durso, R. S. (Ed.), *Handbook of applied cognition* (pp. 551–569). Chichester, England: Wiley.

Mayer, R. E. (1999b). *The promise of educational psychology: Volume 1, Learning in the content areas.* Upper Saddle River, NJ: Prentice Hall.

Mayer, R. E. (2001a). Changing conceptions of learning: A century of progress in the scientific study of education. In L. Corno (Ed.), *Education across a century: The centennial volume, One hundredth yearbook of the National Society for the Study of Education* (pp. 34–75). Chicago: University of Chicago Press.

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

Mayer, R. E. (2002). *The promise of educational psychology: Volume 2, Teaching for meaningful learning.* Upper Saddle River, NJ: Prentice Hall.

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., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93,* 390–397.

Mayer, R. E., & Gallini, J. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82,* 715–726.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93,* 187–198.

Mayer, R. E., & Mathias, A. (2001, April). *Static and interactive pretraining that reduces cognitive load during multimedia learning.* Paper presented at the annual convention of the American Educational Research Association, Seattle, WA.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90,* 312–320.

Mayer, R. E., Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology, 91,* 638–643.

Mayer, R. E., Sims, V., & Tajika, H. (1995). A comparison of how textbooks teach mathematical problem solving in Japan and the United States. *American Educational Research Journal, 32,* 443–460.

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, 84,* 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,* 31–43.

Mayer, R. E., & Wittrock, M.C. (1996). Problem-solving transfer. In D. Berliner and R. Calfee (Eds.), *Handbook of educational psychology* (pp. 45–61). New York: Macmillan.

Meyer, B. J. F. (1975). *The organization of prose and its effects on memory.* New York: Elsevier.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory.* Cambridge, UK: Cambridge University Press.

Mohr, P., Glover, J., & Ronning, R. R. (1984). The effect of related and unrelated details on the recall of major ideas in prose. *Journal of Reading Behavior, 16,* 97–109.

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

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

Moreno, R., & Mayer, R. E. (2000a). 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. (2000b). Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psychology, 93,* 724–733.

Moreno, R., & Mayer, R. E. (2001, August). *Virtual reality and learning: Cognitive and motivational effects of students' sense of presence.* Paper presented at the Ninth International Conference on Human Computer Interaction, New Orleans, LA.

Moreno, R., & Mayer, R. E. (in press). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94,* 000–000.

Moreno, R., Mayer, R. E., & Lester, J. C. (2000). Life-like pedagogical agents in constructivist multimedia environments: Cognitive consequences of their interaction. *ED-MEDIA 2000 Proceedings* (pp. 741–746). Charlottesville, VA: AACE Press.

Moreno, R., Mayer, R. E., Spires, H., & Lester, J. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19,* 177–214.

Mousavi, S., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87,* 319–334.

Norman, D. A. (1993). *Things that make us smart.* Reading, MA: Addison-Wesley.

O'Neil, H. F., Mayer, R. E., Herl, H., Niemi, C., Olin, K., & Thurman, R. A. (2000). Instructional strategies for virtual environments. In H. F. O'Neil & D. H. Andrews (Eds.), *Aircraft training: Methods, technologies, and assessment* (pp. 105–130). Mahwah, NJ: Erlbaum.

Paas, F. G. W. C., & Van Merrienboer, J. J. G. (1994). Measurement of cognitive load in instructional research. *Perceptual & Motor Skills, 79,* 419–430.

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

Piaget, J. (1954). *The construction of reality in the child.* New York: Basic Books.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second language multimedia learning environment. *Journal of Educational Psychology, 90,* 25–36.

Reder, L. M., & Anderson, J. R. (1980). A comparison of texts and their summaries: Memorial consequences. *Journal of Verbal Learning & Verbal Behavior, 19,* 121–134.

Reeves, B., & Nass, C. (1996). *The media equation.* New York: Cambridge University Press.

Renninger, K. A., Hidi, S., & A. Krapp, A. (Eds.). (1992). *The role of interest in learning and development.* Hillsdale, NJ: Erlbaum.

Rickards, J. P., Fajen, B. R., Sullivan, J. F., & Gillespie, G. (1997). Signaling, notetaking, and field independence-dependence in text comprehension and recall. *Journal of Educational Psychology, 89,* 508–517.

Sadoski, M., & Paivio, A. (2001). *Imagery and text: A dual coding theory of reading and writing.* Mahwah, NJ: Erlbaum.

Schnotz, W., & Kulhavy, R. W. (Eds.). (1994). *Comprehension of graphics.* Amsterdam: North-Holland.

Shirey, L. (1992). Importance, interest, and selective attention. In K. A. Renninger, S. Hidi & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 255–277). Hillsdale, NJ: Erlbaum.

Shirey, L., & Reynolds, R. (1988). Effect of interest on attention and learning. *Journal of Educational Psychology, 80,* 159–166.

Sweller, J. (1999). *Instructional design in technical areas.* Camberwell, Australia: ACER Press.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12,* 185–233.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *Journal of Experimental Psychology: General, 119,* 176–192.

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied, 3,* 257–287.

Van Merrienboer, J. J. G. (1997). *Training complex cognitive skills.* Englewood Cliffs, NJ: Educational Technology Publishers.

Wade, S. (1992). How interest affects learning from text. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 255–277). Hillsdale, NJ: Erlbaum.

Wade, S., & Adams, R. (1990). Effects of importance and interest on recall of biographical text. *Journal of Reading Behavior, 22,* 331–353.

Williams, M. D. (1996). Learner control and instructional technologies. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 957–983). New York: Macmillan.

Willows, D. M., & Houghton, H. A. (Ed.). (1987). *The psychology of illustration: Volume 1, Basic research.* New York: Springer-Verlag.

Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist, 24,* 345–376.