Pergamon

Learning and Instruction 13 (2003) 125–139

Learning and Instruction

www.elsevier.com/locate/learninstruc

The promise of multimedia learning: using the same instructional design methods across different media

Richard E. Mayer *

Department of Psychology, University of California, Santa Barbara, CA 93106, USA

Abstract

Multimedia learning occurs when students build mental representations from words and pictures that are presented to them (e.g., printed text and illustrations or narration and animation). The promise of multimedia learning is that students can learn more deeply from well-designed multimedia messages consisting of words and pictures than from more traditional modes of communication involving words alone. This article explores a program of research aimed at determining (a) research-based principles for the design of multimedia explanations—which can be called methods, and (b) the extent to which methods are effective across different learning environments—which can be called media. A review of research on the design of multimedia explanations conducted in our lab at Santa Barbara shows (a) a multimedia effect—in which students learn more deeply from words and pictures than from words alone—in both book-based and computer-based environments, (b) a coherence effect—in which students learn more deeply when extraneous material is excluded rather than included—in both book-based and computer-based environments, (c) a spatial contiguity effect—in which students learn more deeply when printed words are placed near rather than far from corresponding pictures—in both book-based and computer-based environments, and (d) a personalization effect—in which students learn more deeply when words are presented in conversational rather than formal style—both in computer-based environments containing spoken words and those using printed words. Overall, our results provide four examples in which the same instructional design methods are effective across different media.

Keywords: Computer-based learning (CBL); Computer-based instruction (CBI); On-line training; Multimedia learning

* Tel.: +1-805-893-2472; fax: +1-805-893-4303.
E-mail address: mayer@psych.ucsb.edu (R.E. Mayer).

0959-4752/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0959-4752(02)00016-6

1. Introduction

Consider the following learning scenario. A student sits at her desk and listens to a teacher giving a scientific explanation, such as how a bicycle tire pump works. For example, the teacher says: “When the handle is pulled up, the piston moves up, the inlet valve opens, the outlet valve closes, and air enters the cylinder. When the handle is pushed down, the piston moves down, the inlet valve closes, the outlet valves opens, and air moves out through the hose”. Alternatively, a student sits at her desk, opens her book, and reads a passage, such as the explanation of how a bicycle tire pump works. These are single-medium presentations that involve only one way of presenting information—words.

What’s wrong with this verbal-only method of instruction? On the positive side, verbal modes of instruction have a long history in education and words are clearly the dominant vehicle for delivering information in schools. In addition, the lecture and the textbook passage clearly present the key information describing how the pump works. On the negative side, however, verbal modes of instruction are sometimes based on an inadequate conception of how students learn—which can be called the information delivery view. According to this view learning involves adding new information to memory, so teaching involves delivering the information to the learner such as through words. This view is inconsistent with current theories of how people learn (Bransford, Brown, & Cocking, 1999; Bruer, 1993; Lambert & McCombs, 1998; Mayer, Heiser, & Lonn, 2001), namely the constructivist view in which students attempt to make sense of the presented material.

There is also empirical evidence that the verbal-only method does not always work so well. Our research shows, on average, that students who listen to (or read) explanations that are presented solely as words are unable to remember most of the key ideas and experience difficulty in using what was presented to solve new problems (Mayer, 1997, 1999a, 1999b, 2001).

In contrast, consider a learning scenario that goes beyond the purely verbal. A student sits at a computer screen, calls up an on-line encyclopedia, clicks on the entry for “pump”, and views a narrated animation that explains how a bicycle tire pump works. Selected frames from the presentation are shown in Fig. 1, along with corresponding narration indicated in quotation marks. As an alternative, a student may read a book consisting of captioned illustrations; the book shows a series of frames of the pump depicting the steps in the operation of the pump with words that describe each step printed within each frame. These are examples of multimedia learning because the student receives an instructional message that is presented in two formats—as words (spoken or printed text) and pictures (animation or illustrations). Certainly, adding pictures to words does not always improve learning; thus, our goal is to determine the conditions under which adding pictures fosters deep learning.

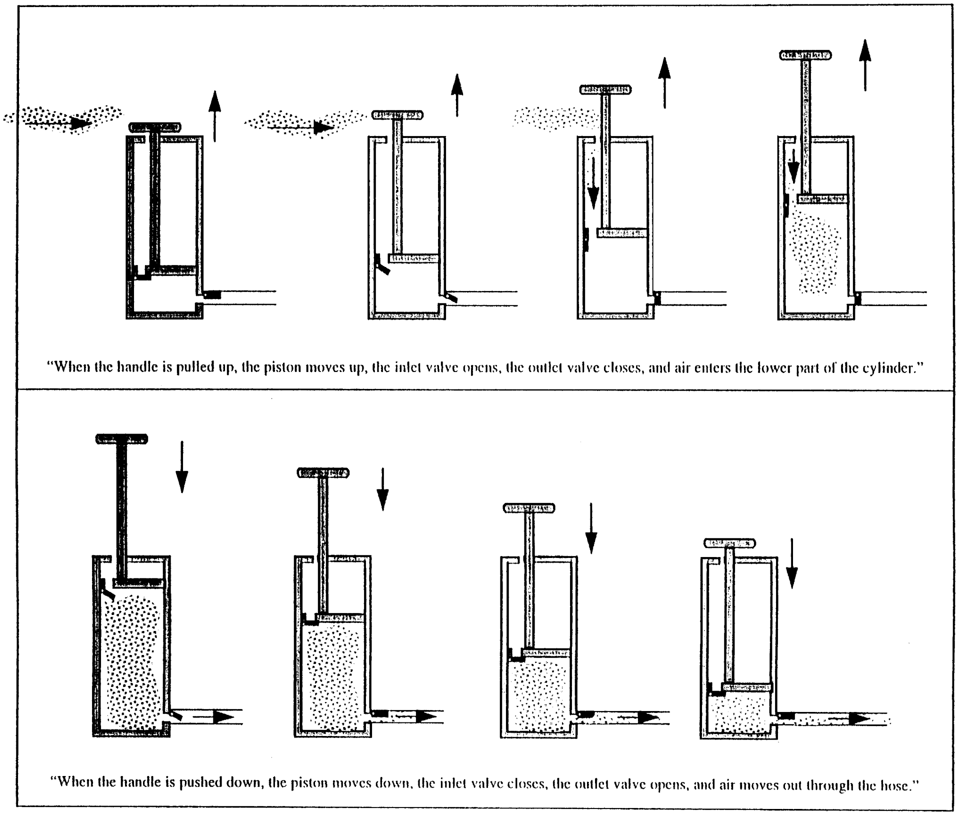

Fig. 1. Selected frames from a multimedia explanation of how a pump works.

2. What is the promise of multimedia learning?

The promise of multimedia learning is that, by combining pictures with words, we will be able to foster deeper learning in students. First, multimedia instruction messages can be designed in ways that are consistent with how people learn, and thus can serve as aids to human learning (Mayer, 1997, 1999a, 1999b, 2001). Second, there is a growing research base showing that students learn more deeply from well-designed multimedia presentations than from traditional verbal-only messages, including improved performance on tests of problem-solving transfer (Mandl & Levin, 1989; Mayer, 2001; Najjar, 1998; Schnotz & Kulhavy, 1994; Sweller, 1999; Van Merrienboer, 1997). In short, the promise of multimedia learning is that teachers can tap the power of visual and verbal forms of expression in the service of promoting student understanding.

For purposes of our research program we define deep learning (or understanding) as learning that leads to problem-solving transfer. For example, someone who understands how pumps work is able to troubleshoot a malfunctioning pump by generating possible reasons for why a pump does not work or is able to tell how to redesign a pump to meet a new purpose (such as making it more reliable).

In order to fulfill the promise of multimedia learning, we need a research base and theoretical framework that provide answers to basic questions. Do students learn more deeply from multimedia messages than from verbal-only ones? Under what conditions does it help to add pictures to words? How does multimedia learning work? Can students engage in active learning when they learn from media that do not allow for much hands-on activity such as multimedia messages? What is the role of technology in promoting learning? Do methods (i.e., design principles for multimedia presentations) work the same way across various media (e.g., book-based or computer-based environments)? These are the kinds of questions that researchers address in the young but growing field of multimedia learning (Mandl & Levin, 1989; Mayer, 2001; Schnotz & Kulhavy, 1994; Sweller, 1999; Van Merrienboer, 1997) and that I explore in this paper.

3. What is a multimedia instructional message?

Let's begin with a very basic and focused definition of what I mean by a multimedia instructional message. A multimedia instructional message is a presentation consisting of words and pictures that is designed to foster meaningful learning. Thus, there are two parts to the definition: (a) the presentation contains words and pictures, and (b) the presentation is designed to foster meaningful learning. As you can see, I have limited my definition of multimedia instructional messages to presentations involving two formats—words and pictures. The words can include printed or spoken text and the pictures can include static graphics (such as illustrations, maps, charts, and photos) and dynamic graphics (such as animation and video). Although broader definitions of multimedia could include modalities such as smell or touch, or formats such as music or non-speech sound, I have chosen to focus on multimedia instructional messages that use words and pictures.

As you can also see, the goal of multimedia instructional messages is to foster meaningful learning. Meaningful learning can be assessed through problem-solving transfer tests in which the learner is asked to use the presented material in new ways. For example, we can ask students to troubleshoot a pump ("Suppose you push down and pull up on the handle of a pump several times, but no air comes out. What could have gone wrong?") or to redesign the pump ("What could be done to make a pump more reliable, that is, to make sure it would not fail?") or to explain the pump's behavior ("Why does air enter the pump? Why does air exit from the pump?"). In addition, the presented material itself must be potentially meaningful—such as an explanation of how something works. According to this definition, multimedia instructional messages would not include the teaching of an arbitrary list of facts or an arbitrary procedure. Thus, I have chosen to focus on multimedia messages that provide explanations of how something works.

4. How does multimedia learning work?

The promise of multimedia learning—that is, promoting student understanding by mixing words and pictures—depends on designing multimedia instructional messages in ways that are consistent with how people learn. In this section, I present a cognitive theory of multimedia learning that is based on three assumptions suggested by cognitive science research about the nature of human learning—the dual channel assumption, the limited capacity assumption, and the active learning assumption.

The dual channel assumption is that humans possess separate information processing systems for visual and verbal representations, and is derived from the research of Paivio (1986; Clark & Paivio, 1991) and Baddeley (1992, 1998). For example, animations are processed in the visual/pictorial channel and spoken words (i.e., narrations) are processed in the auditory/verbal channel. The limited capacity assumption is that the amount of processing that can take place within each information processing channel is extremely limited (Baddeley, 1992, 1998; Chandler & Sweller, 1991; Sweller, 1999; Van Merrienboer, 1997). For example, learners may be able to mentally activate only about a sentence of the narration and about 10 seconds of the animation at any one time. The active learning assumption is that meaningful learning occurs when learners engage in active cognitive processing including paying attention to relevant incoming words and pictures, mentally organizing them into coherent verbal and pictorial representations, and mentally integrating verbal and pictorial representations with each other and with prior knowledge (Mayer, 1999a, 1999b, 2001; Wittrock, 1989). This process of active learning results in a meaningful learning outcome that can support problem-solving transfer.

A framework for the cognitive theory of multimedia learning is presented in Fig. 2. In a computer-based environment, the external representations may include spoken words, which enter through the ears, and animations, which enter through the eyes. The learner must select relevant aspects of the sounds and images for further processing. In addition, the learner may convert some of the spoken words into verbal representations for further processing in the verbal channel whereas some of the animation can be converted into visual representations for further processing in the visual channel. In a book-based environment, the external representations may include printed words and illustrations, both of which initially enter through the eyes. The learner must select relevant aspects of the incoming images for further pro-

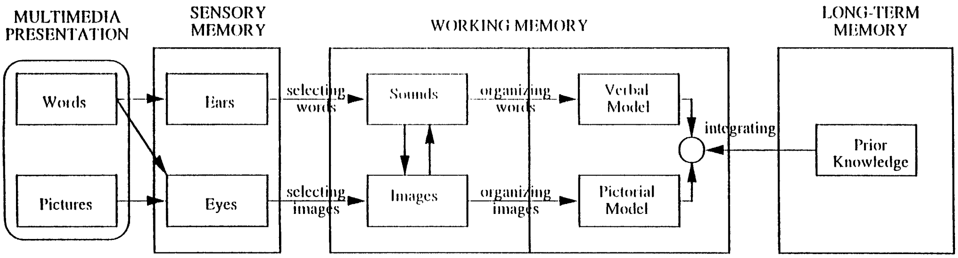

Fig. 2. A framework for a cognitive theory of multimedia learning.

cessing. In addition, the learner may convert some of the printed words into verbal representations to be processed in the verbal channel and may even convert some of the illustrations into verbal representations to be processed in the verbal channel. I refer to these processes as *selecting*.

The second set of processes is to build a coherent mental representation of the verbal material (i.e., form a verbal model) and a coherent mental representation of the visual material (i.e., form a pictorial model)—processes I call *organizing*. A third process is to build connections between the verbal and pictorial models and with prior knowledge—processes I call *integrating*. The processes of selecting, organizing, and integrating generally do not occur in a rigid linear order, but rather in an iterative fashion. Once a learning outcome has been constructed, it is stored in long-term memory for future use. When active learning occurs, the outcome is indexed in long-term memory in a way that allows the learner to use it to solve transfer problems. This model is described more fully in Mayer (2001).

According to the cognitive theory of multimedia learning, meaningful learning depends on all three of these processes occurring for the visual and verbal representations. Instructional methods that enable and promote these processes are more likely to lead to meaningful learning than instructional methods that do not. In this paper, I explore four such methods—presenting words and pictures rather than words alone, excluding extraneous words and pictures, placing corresponding words and pictures near each other on the page or screen, and expressing the words in a conversational style.

According to this view, learners can engage in active learning (such as the processes of selecting, organizing, and integrating) even when the presentation media do not allow hands-on activity (such as printed text and illustrations, or animation and narration). The challenge of multimedia instructional design is to prime and guide active cognitive processing in learners so that learners construct meaningful internal representations. The instructional design methods I explore in this article are intended to achieve the goal of active learning from non-interactive media.

5. Do methods work across media?

Is multimedia learning affected mainly by the instructional method (e.g., well-designed versus less well-designed), by the media (e.g., book-based versus computer-based), or does the effect of the instructional method depend on the media (e.g., well-designed presentations work in one medium but not the other)? There is consensus among instructional design researchers for the value of research on how to design instructional messages across various media (Fleming & Levie, 1993; Najjar, 1998; Van Merrienboer, 1997). Examples of instructional principles include integrating pictures with words, excluding irrelevant material, and presenting words in a conversational style.

In contrast, there is consensus among media researchers that examining the effects of media per se—such as asking whether learning from computers is better than learning from books—is not a productive research strategy (Clark, 1983, 1994;

Clark & Salomon, 1986; Salomon, 1979/1994). For example, Jonassen, Campbell, and Davidson (1994, p. 31) argue for reframing research questions as "learner-centered rather than media-centered" and Kozma (1994, p. 7) calls for research on "media and the methods that employ them as they interact with the cognitive and social processes by which knowledge is constructed". Similarly, Clark (1994) favors research on what learners learn from various instructional methods used within various media rather than on media per se.

Building on this consensus, I have chosen not to focus on the issue of media—whether one media environment (e.g., learning from computers) is better than another (e.g., learning from books). Instead, I have chosen to focus on the issue of *methods across media*—how to design a multimedia message that promotes understanding and whether design principles that work in one media environment (e.g., learning from books) also work in a different media environment (e.g., learning from computers). It should be noted, however, that media are relevant to the extent that some forms of technology enable instructional methods that are not possible with other media. For example, in this review I do not examine the role of interactivity—which is enabled by computer-based technology to a much greater extent than by book-based technology. Instead, I focus on the role of verbal and visual presentation modes that are enabled by computer-based and book-based media. In this section, I explore four instructional design methods and whether they work across different media environments—namely, the multimedia effect, the coherence effect, the spatial contiguity effect, and the personalization effect.

5.1. Multimedia effect with text-and-illustrations and narration-and-animation

The most basic effect of presentation method concerns whether multimedia presentations are more effective than single medium presentations, and in particular, whether adding pictures to words helps students understand an explanation. The *multimedia effect* refers to the finding that students learn more deeply from a multimedia explanation presented in words and pictures than in words alone. Let's consider how this effect fares under two learning environments—with printed text and illustrations on a page and with spoken text and animation on a screen.

For example, consider a textbook lesson on how brakes work in which the explanation is presented in printed words (words-only group) or in printed words along with accompanying illustrations (words-and-pictures group). In a series of three studies involving brakes (Mayer, 1989, Experiments 1 and 2; Mayer & Gallini, 1990, Experiment 1), the words-and-picture group generated a median of 79% more creative solutions on problem-solving transfer tests than the words-only group, yielding an median effect size of 1.50. When the media consist of printed text and illustrations, there is a strong multimedia effect.

Similarly, consider a computer-based lesson on how brakes work that presents a narration describing the steps in how brakes work (words-only group) or narration along with concurrent animation (words-and-pictures group). In a study involving brakes (Mayer & Anderson, 1992, Experiment 2), the words-and-picture group generated 97% more creative solutions that the words-only group, yielding an effect size of 1.67. When the media consist of spoken text and animation, there is a strong multimedia effect.

A similar pattern was found for an explanation of how pumps work. In a study involving printed words and illustrations in a textbook environment (Mayer & Gallini, 1990, Experiment 2), the words-and-pictures group generated 68% more creative solutions than the words-only group, yielding an effect size of 1.00. Similarly, in two studies involving spoken words and animation in a computer-based environment (Mayer & Anderson, 1991, Experiment 2a; Mayer & Anderson, 1992, Experiment 1), the words-and-pictures group generated a median of 108% more creative solutions than the words-only group, yielding an effect size 2.16.

Overall, there is consistent evidence that the multimedia effect can occur across two different media environments—printed text and illustrations on a page as well as spoken text and animation on a screen. It is important to note that while the words used to explain the operation of the brakes or pump were identical across media, the book-based presentations contained additional information about other topics whereas the computer-based presentations did not. Overall, in finding a multimedia effect in both book-based and computer-based environments, we see our first example of the same method effect across different media.

These results are consistent with the cognitive theory of multimedia learning summarized in Fig. 2. In both book-based and computer-based environments, multimedia messages were more likely to cause learners to use both channels in Fig. 2 whereas words-only messages were more likely to cause learners to use only one channel in Fig. 2. Thus, multimedia messages were more likely than words-only messages to prime all the cognitive processes for active learning.

5.2. Coherence effect with text-and-illustrations and narration-and-animation

Another important effect of presentation method concerns what happens when we add interesting but irrelevant material (which can be called *seductive details*) to a multimedia explanation. The *coherence effect* refers to the finding that students learn more deeply from a multimedia explanation when extraneous material is excluded rather than included. Let's consider how the coherence effect holds up under two different media environments—with printed text and illustrations in a book-based environment and with spoken text and animation in a computer-based environment.

For example, students learned how lightning storms develop by reading a passage containing printed words describing the steps in lightning formation along with captioned illustrations depicting the steps in lightning formation. For some students (embellished presentation group), we interspersed a few additional sentences and/or illustrations concerning interesting but irrelevant information, such as a description of how an athlete was struck by lightning on a football field. For other students (concise presentation group), we did not add any seductive details. In five studies involving lightning formation (Harp & Mayer, 1997, Experiment 1; Harp & Mayer, 1998, Experiments 1, 2, 3, and 4), the concise presentation group generated a median of 90% more creative solutions than the embellished presentation group, yielding a median effect size of 1.67. Clearly, there is a strong coherence effect for this book-based multimedia environment.

We also examined how students learned about lightning formation by viewing an animation depicting the steps in the lightning formation along with a concurrent narration describing the steps in lightning formation. For some students (embellished presentation group), we interspersed a few sentences and/or video clips concerning interesting but irrelevant information, similar to the seductive details used in the book-based environment. Other students (concise presentation group) did not receive the additional narration and/or video clips. In three tests involving lightning formation (Mayer, Heiser, & Lonn, 2001), the concise presentation group generated a median of 34% more creative solutions than the embellished presentation group, yielding a median effect size of 0.59. As with the book-based environment, the results across all three studies involving a computer-based environment produced a coherence effect.

Overall, we found consistent evidence for a coherence effect across two different media environments—adding extraneous material hurt student understanding when the multimedia explanation was presented as printed words and illustrations, as well as when it was presented as spoken words and narration. Again, as with the multimedia effect, the coherence effect provides a second example of finding the same method effect across different media.

The coherence effect is consistent with the cognitive theory of multimedia learning summarized in Fig. 2. Adding extraneous words or pictures to a multimedia message can interfere with any of the cognitive processes shown in the figure, by encouraging learners to pay attention to words or images that are not relevant, by disrupting how learners organize words or pictures into a causal chain, and by priming inappropriate schemas to be used to assimilate the incoming words and pictures. In some cases, of course, the goal of instruction may be to consider many alternative perspectives, in which case coherence should not be a guiding instructional design principle.

5.3. Contiguity effect with text-and-illustrations and text-and-animation

Another presentation method effect concerns how we present corresponding text and illustrations on a page or computer screen. The *spatial contiguity effect* is that students learn more deeply from multimedia explanations when corresponding words and pictures are presented near to rather than far from each other on the page or screen. We examined the spatial contiguity effect in two different media environments—with printed text and illustrations in a book-based environment and with printed text and animation in a computer-based environment.

In a book-based environment, some students (integrated presentation group) read a passage on lightning formation, consisting of illustrations that depicted the steps in lightning formation along with corresponding captions that described the steps in words. Other students (separated presentation group) read the identical words and saw the identical illustrations but the corresponding words and pictures were on different pages. In three studies involving lightning formation (Mayer, Steinhoff, Bower, & Mars, 1995, Experiments 1, 2, and 3), students in the integrated presentation group generated a median of 78% more creative solutions to problem-solving questions than did students in the separated presentation group, yielding a median effect size of 1.12. These results demonstrate a spatial contiguity effect within the context of a book-based environment.

In order to determine whether spatial contiguity effects also occurred in computer-based environments, we asked students to view an animation depicting the steps in lightning formation along with concurrent on-screen text that described the steps in words. For some students (integrated presentation group), we placed the on-screen text next to the event that it described; for example, if the on-screen text said, "The negatively charged particles fall to the bottom of the cloud, and most of the positively charged particles rise to the top", that statement was printed next to the cloud on the screen. For other students (separated presentation group), we placed the on-screen text at the bottom of the screen far from the center of action in the animation. In the one study that we conducted using a lightning explanation (Moreno & Mayer, 1999, Experiment 1), students in the integrated presentation group generated 43% more creative solutions on the problem-solving transfer test than did students in the separated presentation group. The corresponding effect size was 0.48. As with the book-based environment, there was a spatial contiguity effect in the computer-based environment.

The same spatial contiguity effect was obtained in both media environments even though the spatial contiguity manipulation—i.e., the physical space between corresponding words and pictures—was not as great in the computer-based environment as in the book-based environment. These results provide the third example of presentation methods having the same kinds of effects across different media.

Like the foregoing effects, spatial contiguity effects are consistent with the cognitive theory of multimedia learning summarized in Fig. 2. When corresponding words and pictures are presented near each other, learners are more likely to be able to hold corresponding words and pictures in working memory at the same time. This situation enables the process of integrating visual and verbal models, a key step in active learning. Of course, it must be noted that external representations (such as text and illustrations) do not translate automatically into internal representations (such as verbal and visual models). That translation process depends on internal cognitive processing—including selecting and organizing—that are subject to working memory limitations. Ainsworth, Bibby, and Wood (1998) have shown that under some circumstances, instruction with multiple representations can be less effective than less demanding instruction.

5.4. Personalization effect with animation-and-narration and animation-and-text

As our fourth and final example, let's consider what happens when we personalize a multimedia presentation by changing the text from formal to conversational style. The *personalization effect* is that students learn more deeply from a multimedia explanation when the words are presented in conversational style rather than formal style. We examined the personalization effect in two different computer-based environments—animation with narration versus animation with on-screen text. In both cases students viewed an animation, but in one environment the corresponding words were presented as spoken text and in the other environment the corresponding words were presented as printed text placed at the bottom of the screen.

In the animation-and-narration environment, students viewed an animation depicting the steps in the formation of lightning while they listened to concurrent narration that described the steps in lightning formation. For some students (conventional presentation group) the wording was formal, and was similar to the third-person monologue style used in many textbooks and lectures. The script was the same as that used in other studies involving lightning as described in previous sections. For other students (personalized presentation), some of the words were changed into first and second person (e.g., saying "your cloud" instead of "the cloud") and some conversational sentences were added (e.g., "As you watch, you tilt your head skyward"). However, both versions contained the same explanation based on the same steps in lightning formation. In one study involving animation and narration (Moreno & Mayer, 2000, Experiment 1), the personalized group generated 36% more creative solutions than the conventional group on the problem-solving transfer test, yielding an effect size of 0.96. When the media include animation and narration, there is evidence for a personalization effect.

What happens to the personalization effect when the words are presented as printed text rather than spoken text? In the animation-and-text environment, students viewed an animation depicting the steps in lightning formation along with concurrent on-screen text presented at the bottom of the screen as one or two sentences at a time. For some students (conventional presentation group), the wording was identical to the formal style used in the animation-and-narration environment. For other students (personalized presentation group), the wording was identical to the conversational style used in the animation-and-narration environment. In one study (Moreno & Mayer, 2000, Experiment 2), the personalized group generated 116% more creative solutions than the conventional group on the problem-solving transfer test. The corresponding effect size was 1.60. Clearly, when the media include animation and on-screen text, there is evidence for a personalization effect.

Overall, a personalization effect was obtained across two different media—when the explanation was presented as animation and narration and when the explanation was presented as animation and on-screen text. The results are preliminary because only one study was involved for each of the media environments, but the consistent pattern produces our fourth and final example of the same method effects with different media.

The personalization effect is consistent with an expanded version of the cognitive theory of multimedia learning shown in Fig. 2. Personalized messages may prime the conversation schema in learners—that is, learners may be more willing to accept that they are in a human-to-human conversation including all the conventions of trying hard to understand what the other person is saying. In trying harder to understand the message, learners with personalized messages are more likely to engage in the cognitive processes shown in the figure, particularly organizing and integrating.

6. Conclusion

As summarized in Table 1, our results provide four case examples of a straightforward finding: instructional design methods that promote deeper learning in one media environment (such as text and illustrations) also promote deep learning in other media environments (such as narration and animation). This means that good instructional methods can work across media. In short, the principles of instructional design do not necessarily change when the learning environment changes.

There may be some aspects of good instructional design that are unique to a particular medium, perhaps because that medium enables instructional methods not available elsewhere. Similarly, some aspects of good instructional design are unique to the to-be-taught material. We selected materials—explanations of how things work—that depended on both words and pictures whereas some materials may require only words.

Why do our instructional methods work across media? In taking a learner-centered approach, I focus on the nature of human learning. The human information processing system—as represented in Fig. 2—remains constant across different media environments. Humans possess two channels—visual and verbal—regardless of whether material is presented by book or by computer. Each channel is limited in capacity regardless of whether material is presented by book or computer. Active cognitive processing—including selecting, organizing, and integrating mental representations—promotes meaningful learning regardless of whether material is presented by book or computer.

A basic requirement in multimedia learning situations is that learners be able to

Table 1
Effect size (and number of studies) for four examples of the same method effects across different media environments

Method effect (and lesson content)	Media environment	
	Text and illustrations[a]	Narration and animation
	Effect size (n)	Effect size (n)
Multimedia effect		
(brakes)	1.50 (3)	1.67 (1)
(pumps)	1.00 (1)	2.16 (2)
Coherence effect		
(lightning)	1.67 (5)	0.59 (3)
Spatial contiguity effect		
(lightning)	1.12 (3)	0.48 (1)
Personalization effect		
(lightning)	1.60 (1)	0.96 (1)

[a] For personalization effect the first column involves text and animation; n refers to number of studies that effect size is based on.

hold corresponding visual and verbal representations in working memory at the same time (such as the image of the inlet valve opening and the proposition, "the inlet valve opens"). Instructional methods that enable this condition are likely to be effective across media environments. In rejecting a technology-centered approach, I conclude that media environments do not cause learning, cognitive processing by the learner causes learning. If an instructional method promotes the same kinds of cognitive processing across different media, then it will result in the same benefits across media.

Each of the four methods I have summarized in this article was intended to promote active cognitive processing within the constraints of the human information processing system: adding pictures to words, eliminating extraneous words and pictures, placing words near corresponding pictures, and using conversational style for words. In short, the design of multimedia instructional messages should be based on an understanding of the nature of human learning. Using different technologies does not change the fundamental nature of how the human mind works; however, to the extent that instructional technologies are intelligently designed, they can serve as powerful aids to human cognition.

In our research, we were able to significantly increase how well students understood scientific explanations by basing multimedia message design on an understanding of how the human mind works. Understanding was measured by tests of problem-solving transfer, in which students were asked to generate as many useful solutions as possible to problems they had not seen before. Redesigning multimedia explanations to mesh with the way humans learn enabled students to generate more creative solutions to problem-solving transfer questions—ranging from gains of 36% to 116%. As shown in Table 1, effect sizes ranged from 0.48 to 2.16, with most averaging above 1.00. Although our instructional materials were modest—a short lesson about a single scientific system—the results consistently demonstrate the benefits of learner-centered design.

In conclusion, the promise of multimedia learning rests in the potential of using words and pictures in ways that promote meaningful learning. The design principles summarized in this article and elsewhere (Mayer, 1997, 1999a, 1999b, 2001) are examples of multimedia instructional design based on a cognitive theory of multimedia learning. My premise in this research is that the intelligent design of multimedia instructional messages depends on an understanding of how the human mind works. The role of instructional technology in this learner-centered scenario is to serve as a tool that increases the power of human cognition. In our research this increase in the power of human cognition is measured in terms of improvements in problem-solving transfer. To the extent that multimedia researchers remain focused on the learner-centered issue of how to promote appropriate cognitive processing in learners, the promise of multimedia learning remains strong.

Acknowledgements

Some of the material in this article was presented as a keynote address entitled "The promise of multimedia learning", at the 8th European Conference for Research on Learning and Instruction in Goteborg, Sweden, August 1999. I gratefully acknowledge colleagues with whom I have worked on research in multimedia learning: Richard B. Anderson, Joan Gallini, Shannon Harp, and Roxana Moreno.

References

Ainsworth, S. E., Bibby, P. A., & Wood, D. J. (1998). Analysing the costs and benefits of multi-representational learning environments. In M. van Someren, P. Reimann, H. P. Boshuizen, & T. de Jong (Eds.), *Learning with multiple representations* (pp. 120–136). Oxford, UK: Pergamon Press.

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

Baddeley, A. (1998). *Human memory*. Boston, MA: Allyn and Bacon.

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

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

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

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

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, *53*, 445–459.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, *42*, 21–30.

Clark, R. E., & Salomon, G. (1986). Media in teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.) (pp. 464–478). New York: Macmillan.

Fleming, M., & Levie, W.H. (Eds.). (1993). *Instructional message design: Principles from the behavioral and cognitive sciences* (2nd ed.). Englewood Cliffs, NJ: Educational Technology Publications.

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.

Jonassen, D. H., Campbell, J. P., & Davidson, M. E. (1994). Learning with media: restructuring the debate. *Educational Technology Research and Development*, *42*, 31–39.

Kozma, R. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development*, *42*, 1–19.

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

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

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

Mayer, R. E. (1997). Multimedia learning: are we asking the right questions? *Educational Psychologist*, *32*, 1–19.

Mayer, R. E. (1999a). Multimedia aids to problem-solving transfer. *International Journal of Educational Research*, *31*, 611–623.

Mayer, R. E. (1999b). Research-based principles for the design of instructional messages. *Document Design*, *1*, 7–20.

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

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., & 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., 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.

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

Moreno, R., & Mayer, R. E. (2000). Engaging students in active learning: the case for personalized multimedia messages. *Journal of Educational Psychology*, *92*, 724–733.

Najjar, L. J. (1998). Principles of educational multimedia user interface design. *Human Factors*, *40*, 311–323.

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

Salomon, G. (1979/1994). *Interaction of media, cognition, and learning*. Hillsdale, NJ: Erlbaum.

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

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

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

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