EDUCATIONAL PSYCHOLOGIST, *38*(1), 43–52

Nine Ways to Reduce Cognitive Load in Multimedia Learning

Richard E. Mayer

Department of Psychology
University of California, Santa Barbara

Roxana Moreno

Educational Psychology Program
University of New Mexico

First, we propose a theory of multimedia learning based on the assumptions that humans possess separate systems for processing pictorial and verbal material (dual-channel assumption), each channel is limited in the amount of material that can be processed at one time (limited-capacity assumption), and meaningful learning involves cognitive processing including building connections between pictorial and verbal representations (active-processing assumption). Second, based on the cognitive theory of multimedia learning, we examine the concept of cognitive overload in which the learner's intended cognitive processing exceeds the learner's available cognitive capacity. Third, we examine five overload scenarios. For each overload scenario, we offer one or two theory-based suggestions for reducing cognitive load, and we summarize our research results aimed at testing the effectiveness of each suggestion. Overall, our analysis shows that cognitive load is a central consideration in the design of multimedia instruction.

WHAT IS MULTIMEDIA LEARNING AND INSTRUCTION?

The goal of our research is to figure out how to use words and pictures to foster meaningful learning. We define *multimedia learning* as learning from words and pictures, and we define *multimedia instruction* as presenting words and pictures that are intended to foster learning. The words can be printed (e.g., on-screen text) or spoken (e.g., narration). The pictures can be static (e.g., illustrations, graphs, charts, photos, or maps) or dynamic (e.g., animation, video, or interactive illustrations). An important example of multimedia instruction is a computer-based narrated animation that explains how a causal system works (e.g., how pumps work, how a car's braking system works, how the human respiratory system works, how lightning storms develop, how airplanes achieve lift, or how plants grow).

We define *meaningful learning* as deep understanding of the material, which includes attending to important aspects of the presented material, mentally organizing it into a coherent cognitive structure, and integrating it with relevant existing knowledge. Meaningful learning is reflected in the ability to apply what was taught to new situations, so we measure learning outcomes by using problem-solving transfer tests (Mayer & Wittrock, 1996). In our research, meaningful learning involves the construction of a mental model of how a causal system works. In addition to asking whether learners can recall what was presented in a lesson (i.e., retention test), we also ask them to solve novel problems using the presented material (i.e., transfer test). All the results reported in this article are based on problem-solving transfer performance.

In pursuing our research on multimedia learning, we have repeatedly faced the challenge of cognitive load: Meaningful learning requires that the learner engage in substantial cognitive processing during learning, but the learner's capacity for cognitive processing is severely limited. Instructional designers have come to recognize the need for multimedia instruction that is sensitive to cognitive load (Clark, 1999; Sweller, 1999; van Merriënboer, 1997). A central challenge facing designers of multimedia instruction is the potential for cognitive overload—in which the learner's intended cognitive processing exceeds the learner's available cognitive capacity. In this article we present a theory of how people learn from multimedia instruction, which highlights the potential for cognitive overload. Then, we describe how to design multimedia in-

Requests for reprints should be sent to Richard E. Mayer, Department of Psychology, University of California, Santa Barbara, CA 93106–9660. E-mail: mayer@psych.ucsb.edu

struction in ways that reduce the chances of cognitive overload in each of five overload scenarios.

HOW THE MIND WORKS

We begin with three assumptions about how the human mind works based on research in cognitive science—the dual channel assumption, the limited capacity assumption, and the active processing assumption. These assumptions are summarized in Table 1.

TABLE 1
Three Assumptions About How the Mind Works in Multimedia Learning

Assumption	*Definition*
Dual channel	Humans possess separate information processing channels for verbal and visual material.
Limited capacity	There is only a limited amount of processing capacity available in the verbal and visual channels.
Active processing	Learning requires substantial cognitive processing in the verbal and visual channels.

First, the human information-processing system consists of two separate channels—an auditory/verbal channel for processing auditory input and verbal representations and a visual/pictorial channel for processing visual input and pictorial representations.[1] The dual-channel assumption is a central feature of Paivio's (1986) dual-coding theory and Baddeley's (1998) theory of working memory, although all theorists do not characterize the subsystems exactly the same way (Mayer, 2001).

Second, each channel in the human information-processing system has limited capacity—only a limited amount of cognitive processing can take place in the verbal channel at any one time, and only a limited amount of cognitive processing can take place in the visual channel at any one time. This is the central assumption of Chandler and Sweller's (1991; Sweller, 1999) cognitive load theory and Baddeley's (1998) working memory theory.

Third, meaningful learning requires a substantial amount of cognitive processing to take place in the verbal and visual channels. This is the central assumption of Wittrock's (1989) generative-learning theory and Mayer's (1999, 2002) selecting–organizing–integrating theory of active learning. These processes include paying attention to the presented material, mentally organizing the presented material into a coherent structure, and integrating the presented material with existing knowledge.

Let us explore these three assumptions within the context of a cognitive theory of multimedia learning that is summarized in Figure 1. The theory is represented as a series of boxes arranged into two rows and five columns, along with arrows connecting them. The two rows represent the two information-processing channels, with the auditory/verbal channel on top and the visual/pictorial channel on the bottom. This aspect of the Figure 1 is consistent with the dual-channel assumption.

The five columns in Figure 1 represent the modes of knowledge representation—physical representations (e.g., words or pictures that are presented to the learner), sensory representations (in the ears or eyes of the learner), shallow working memory representations (e.g., sounds or images attended to by the learner), deep working memory representations (e.g., verbal and pictorial models constructed by the learner), and long-term memory representations (e.g., the learner's relevant prior knowledge). The capacity for physically presenting words and pictures is virtually unlimited, and the capacity for storing knowledge in long-term memory is virtually unlimited, but the capacity for mentally holding and manipulating words and images in working memory is limited. Thus, the working memory columns in Figure 1 are subject to the limited-capacity assumption.

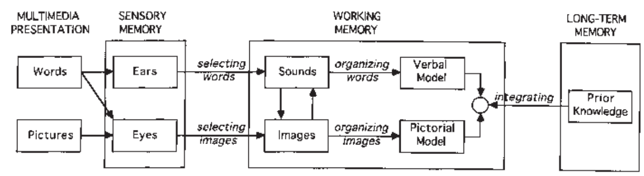

FIGURE 1 Cognitive theory of multimedia learning.

The arrows represent cognitive processing. The arrow from words to eyes represents printed words impinging on the eyes; the arrow from *words* to *ears* represents spoken words impinging on the ears; and the arrow from *pictures* to *eyes* represents pictures (e.g., illustrations, charts, photos, animations, and videos) impinging on the eyes. The arrow labeled *selecting words* represents the learner's paying attention to some of the auditory sensations coming in from the ears, whereas the arrow labeled *selecting images* represents the learner's paying attention to some of the visual sensations coming in through the eyes.[2] The arrow labeled *organizing words* represents the learner's constructing a coherent verbal representation from the incoming words, whereas the arrow labeled *organizing images* represents the learner's constructing a coherent pictorial representation from the incoming images. Finally, the arrow labeled *integrating* represents the merging of the verbal model, the pictorial model, and relevant prior knowledge. In addition, we propose that the selecting

[1] Based on research on discourse processing (Graesser, Millis, & Zwaan, 1997), it is not appropriate to equate a verbal channel with an auditory channel. Mayer (2001) provided an extended discussion of the nature of dual channels.

[2] *Selecting words* refers to selecting aspects of the text information rather than only specific words. *Selecting images* refers to selecting parts of pictures rather than only whole pictures.

and organizing processes may be guided partially by prior knowledge activated by the learner. In multimedia learning, active processing requires five cognitive processes: selecting words, selecting images, organizing words, organizing images, and integrating. Consistent with the active-processing assumption, these processes place demands on the cognitive capacity of the information-processing system. Thus, the labeled arrows in Figure 1 represent the active processing required for multimedia learning.

THE CASE OF COGNITIVE OVERLOAD

Let us consider what happens in multimedia learning, that is, a learning situation in which words and pictures are presented. A potential problem is that the processing demands evoked by the learning task may exceed the processing capacity of the cognitive system—a situation we call *cognitive overload*. The ever-present potential for cognitive overload is a central challenge for instructors (including instructional designers) and learners (including multimedia learners); meaningful learning often requires substantial cognitive processing using a cognitive system that has severe limits on cognitive processing.

We distinguish among three kinds of cognitive demands: essential processing, incidental processing, and representational holding.[3] *Essential processing* refers to cognitive processes that are required for making sense of the presented material, such as the five core processes in the cognitive theory of multimedia learning—selecting words, selecting images, organizing words, organizing images, and integrating. For example, in a narrated animation presented at a fast pace and consisting of unfamiliar material, essential processing involves using a great deal of cognitive capacity in selecting, organizing, and integrating the words and the images.

Incidental processing refers to cognitive processes that are not required for making sense of the presented material but are primed by the design of the learning task. For example, adding background music to a narrated animation may increase the amount of incidental processing to the extent that the learner devotes some cognitive capacity to processing the music.

Representational holding refers to cognitive processes aimed at holding a mental representation in working memory over a period of time. For example, suppose that an illustration is presented in one window and a verbal description of it is presented in another window, but only one window can appear on the screen at one time. In this case, the learner must hold a representation of the illustration in working memory while reading the verbal description or must hold a representation of the verbal information in working memory while viewing the illustration.

Table 2 summarizes the three kinds of cognitive-processing demands in multimedia learning. The total processing intended for learning consists of essential processing plus incidental processing plus representational holding. Cognitive overload occurs when the total intended processing exceeds the learner's cognitive capacity.[4] Reducing cognitive load can involve redistributing essential processing, reducing incidental processing, or reducing representational holding.

In the following sections, we explore nine ways to reduce cognitive load in multimedia learning. We describe five different scenarios involving cognitive overload in multimedia learning. For each overload scenario we offer one or two suggestions regarding how to reduce cognitive overload based on the cognitive theory of multimedia learning, and we review the effectiveness of our suggestions based on a 12-year program of research carried out at the University of California, Santa Barbara (UCSB). Our recommendations for reducing cognitive load in multimedia learning are summarized in Table 3.

Type 1 Overload: Off-Loading When One Channel is Overloaded With Essential Processing Demands

Problem: One channel is overloaded with essential processing demands. Consider the following situation: A student is interested in understanding how lightning works. She goes to a multimedia encyclopedia and clicks on the entry for *lightning*. On the screen appears a 2-min animation depicting the steps in lightning formation along with concurrent on-screen text describing the steps in lightning formation. The on-screen text is presented at the bottom on the screen, so while the student is reading she cannot view the animation, and while she is viewing the animation she cannot read the text.

This situation creates what Sweller (1999) called a *split-attention effect* because the learner's visual attention is split between viewing the animation and reading the

TABLE 2
Three Kinds of Demands for Cognitive Processing in Multimedia Learning

Type of Processing	*Definition*
Essential processing	Aimed at making sense of the presented material including selecting, organizing, and integrating words and selecting, organizing, and integrating images.
Incidental processing	Aimed at nonessential aspects of the presented material.
Representational holding	Aimed at holding verbal or visual representations in working memory.

[3] *Essential processing* corresponds to the term *germane load* as used in the introduction to this special issue. *Incidental processing* corresponds to the term *extraneous load* as used in the introduction to this special issue. Finally, *representational holding* is roughly equivalent to the term *intrinsic load*.

[4] To maintain conceptual clarity, we use the term *processing demands* to refer to properties of the learning materials or situation and the term *processing* to refer to internal cognitive activity of learners.

TABLE 3
Load-Reduction Methods for Five Overload Scenarios in Multimedia Instruction

Type of Overload Scenario	*Load-Reducing Method*	*Description of Research Effect*	*Effect Size*
Type 1: Essential processing in visual channel > cognitive capacity of visual channel			
Visual channel is overloaded by essential processing demands.	Off-loading: Move some essential processing from visual channel to auditory channel.	Modality effect: Better transfer when words are presented as narration rather than as on-screen text.	1.17 (6)
Type 2: Essential processing (in both channels) > cognitive capacity			
Both channels are overloaded by essential processing demands.	Segmenting: Allow time between successive bite-size segments.	Segmentation effect: Better transfer when lesson is presented in learner-controlled segments rather than as continuous unit.	1.36 (1)
	Pretraining: Provide pretraining in names and characteristics of components.	Pretraining effect: Better transfer when students know names and behaviors of system components.	1.00 (3)
Type 3: Essential processing + incidental processing (caused by extraneous material) > cognitive capacity			
One or both channels overloaded by essential and incidental processing (attributable to extraneous material).	Weeding: Eliminate interesting but extraneous material to reduce processing of extraneous material.	Coherence effect: Better transfer when extraneous material is excluded.	0.90 (5)
	Signaling: Provide cues for how to process the material to reduce processing of extraneous material.	Signaling effect: Better transfer when signals are included.	0.74 (1)
Type 4: Essential processing + incidental processing (caused by confusing presentation) > cognitive capacity			
One or both channels overloaded by essential and incidental processing (attributable to confusing presentation of essential material).	Aligning: Place printed words near corresponding parts of graphics to reduce need for visual scanning.	Spatial contiguity effect: Better transfer when printed words are placed near corresponding parts of graphics.	0.48 (1)
	Eliminating redundancy: Avoid presenting identical streams of printed and spoken words.	Redundancy effect: Better transfer when words are presented as narration rather narration and on-screen text.	0.69 (3)
Type 5: Essential processing + representational holding > cognitive capacity			
One or both channels overloaded by essential processing and representational holding.	Synchronizing: Present narration and corresponding animation simultaneously to minimize need to hold representations in memory.	Temporal contiguity effect: Better transfer when corresponding animation and narration are presented simultaneously rather than successively.	1.30 (8)
	Individualizing: Make sure learners possess skill at holding mental representations.	Spatial ability effect: High spatial learners benefit more from well-designed instruction than do low spatial learners.	1.13 (2)

Note. Numbers in parentheses indicate number of experiments on which effect size was based.

on-screen text. This problem is represented in Figure 1 by the arrow from *picture* to *eyes* (for the animation) and the arrow from *words* to *eyes* (for the on-screen text); thus, the eyes receive a lot of concurrent information, but only some of that information can be selected for further processing in visual working memory (i.e., the arrow from *eyes* to *images* can only carry a limited amount of information).

Solution: Off-loading. One solution to this problem is to present words as narration. In this way, the words are processed—at least initially—in the verbal channel (indicated by the arrow from *words* to *ears* in Figure 1), whereas the animation is processed in the visual channel (indicated by the arrow from *picture* to *eyes* in Figure 1). The processing demands on the visual channel are thereby reduced, so the learner is better able to select important aspects of animation for further processing (indicated by the arrow from *eyes* to *image*). The processing demands on the verbal channel are also moderate, so the learner is better able to select important aspects of the narration for further processing (indicated by the arrow from *ears* to *sounds*). In short, the use of narrated animation represents a method for off-loading (or reassigning) some of the processing demands from the visual channel to the verbal channel.

In a series of six studies carried out in our laboratory at UCSB, students performed better on tests of problem-solving transfer when scientific explanations were presented as animation and narration rather than as animation and on-screen text (Mayer & Moreno, 1998, Experiments 1 and 2; Moreno & Mayer, 1999, Experiments 1 and 2; Moreno, Mayer, Spires, & Lester, 2001, Experiments 4 and 5). The median effect size was 1.17. We refer to this result as a *modality effect*: Students understand a multimedia explanation better when the words are presented as narration rather than as on-screen text. A similar effect was reported by Mousavi, Low, and Sweller (1995) in a book-based multimedia environment.

The robustness of the modality effect provides strong evidence for the viability of off-loading as a method of reducing cognitive load.

Type 2 Overload: Segmenting and Pretraining When Both Channels are Overloaded With Essential Processing Demands in Working Memory

Problem: Both channels are overloaded with essential processing demands. Suppose a student views a narrated animation that explains the process of lightning formation based on the strategies discussed in the previous section. In this case, some of the narration is selected to be processed as words in the verbal channel and some of the animation is selected to be processed as images in the visual channel (as shown by the arrows in Figure 1 labeled *selecting words* and *selecting images*, respectively). However, if the information content is rich and the pace of presentation is fast, learners may not have enough time to engage in the deeper processes of organizing the words into a verbal model, organizing the images into a visual model, and integrating the models (as shown by the *organizing words*, *organizing images*, and *integrating* arrows in Figure 1). By the time the learner selects relevant words and pictures from one segment of the presentation, the next segment begins, thereby cutting short the time needed for deeper processing.

This situation leads to cognitive overload in which available cognitive capacity is not sufficient to meet the required processing demands. Sweller (1999) referred to this situation as one in which the presented material has high-intrinsic load; that is, the material is conceptually complex. Although it might not be possible to simplify the presented material, it is possible to allow learners to digest intellectually one chunk of it before moving on to the next.

Solution: Segmenting. A potential solution to this problem is to allow some time between successive segments of the presentation. In segmenting, the presentation is broken down into bite-size segments. The learner is able to select words and select images from the segment; the learner also has time and capacity to organize and integrate the selected words and images. Then, the learner is ready for the next segment, and so on. In contrast, when the narrated animation is presented continuously—without time breaks between segments—the learner can select words and select images from the first segment; but, before the learner is able to complete the additional processes of organizing and integration, the next segment is presented, which demands the learner's attention for selecting words and images.

For example, Mayer and Chandler (2001, Experiment 2) broke a narrated animation explaining lightning formation into 16 segments. Each segment contained one or two sentences of narration and approximately 8 to 10 sec of animation. After each segment was presented, the learner could start the next segment by clicking on a button labeled *CONTINUE.* Although students in both groups received identical material, the segmented group had more study time. Students who received the segmented presentation performed better on subsequent tests of problem-solving transfer than did students who received a continuous presentation. The effect size in the one study we conducted was 1.36. We refer to this as a *segmentation effect:* Students understand a multimedia explanation better when it is presented in learner-controlled segments rather than as a continuous presentation. Further research is needed to determine the separate effects of segmenting and interactivity, such as comparing how students learn from multimedia presentations that contain built-in or user-controlled breaks after each segment.

Solution: Pretraining. Although segmenting appears to be a promising technique for reducing cognitive load, sometimes segmenting might not be feasible. An alternative technique for reducing cognitive load when both channels are overloaded with essential processing demands is pretraining, in which learners receive prior instruction concerning the components in the to-be-learned system. Constructing a mental model involves two steps—building component models (i.e., representations of how each component works) and building a causal model (i.e., a representation of how a change in one part of the system causes a change in another part, etc.). In processing a narrated animation explaining how a car's braking system works, learners must simultaneously build component models (concerning how a piston can move forward and back, how a brake shoe can move forward or back, etc.) and a causal model (when the piston moves forward, brake fluid is compressed, etc.). By providing pretraining about the components, learners can more effectively process a narrated animation—devoting their cognitive processing to building a causal model. Without pretraining, students must try to understand each component and the causal links between them—a task that can easily overload working memory.

In a series of three studies involving narrated animations about how brakes work and how pumps work, students performed better on problem-solving transfer tests when the narrated animation was preceded by a short pretraining about the names and behavior of the components (Mayer, Mathias, & Wetzell, 2002, Experiments 1, 2, and 3). The median effect size comparing the pretrained and nonpretrained groups was 1.00. Similar results were reported by Pollock, Chandler, and Sweller (2002). We refer to this result as a *pretraining effect:* Students understand a multimedia presentation better when they know the names and behaviors of the components in the system. Pretraining involves a specific sequencing strategy in which components are presented before a causal system is presented. The results provide support for pretraining as a useful method of reducing cognitive load.

Type 3 Overload: Weeding and Signaling When the System is Overloaded by Incidental Processing Demands Due to Extraneous Material

Problem: One or both channels are overloaded by the combination of essential and incidental processing demands. In the two foregoing scenarios, the cognitive system was required to engage in too much essential processing—such as when complex material is presented at a fast rate. Let us consider a somewhat different overload scenario in which a learner seeks to engage in both essential and incidental processing, which together exceed the learner's available cognitive capacity. For example, suppose a learner clicks on the entry for *lightning* in a multimedia encyclopedia, and he or she receives a narrated animation describing the steps in lightning formation (which requires essential processing) along with background music or inserted narrated video clips of damage caused by lightning (which requires incidental processing).

According to the cognitive theory of multimedia learning, adding interesting but extraneous[5] material to a narrated animation may cause the learner to use limited cognitive resources on incidental processing, leaving less cognitive capacity for essential processing. As a result, the learner will be less likely to engage in the cognitive processes required for meaningful learning of how lightning works—indicated by the arrows in Figure 1. Sweller (1999) referred to the addition of extraneous material in an instructional presentation as an example of extraneous load.

Solution: Weeding. To solve this problem, we suggest eliminating interesting but extraneous material—a load-reducing technique can be called *weeding*. Weeding involves making the narrated animation as concise and coherent as possible, so the learner will not be primed to engage in incidental processing. In a concise narrated animation, the learner is primed to engage in essential processing. In contrast, in an embellished narrated animation—such as one containing background music or inserted narrated video of lightning damage—the learner is primed to engage in both essential and incidental processing.

In a series of five studies carried out in our laboratory at UCSB, students performed better on problem-solving transfer tests after receiving a concise narrated animation than an embellished narrated animation (Mayer, Heiser, & Lonn, 2001, Experiments 1, 3, and 4; Moreno & Mayer, 2000, Experiments 1 and 2). The added material in the embellished narrated animation consisted of background music or adding short narrated video clips showing irrelevant material. The median effect size was .90. We refer to this result as a *coherence effect:* Students understand a multimedia explanation better when interesting but extraneous material is excluded rather than included. The robustness of the coherence effect provides strong evidence for the viability of weeding as a method for reducing cognitive load. Weeding seems to help facilitate the process of selecting relevant information.

Solution: Signaling. When it is not feasible to remove all the embellishments in a multimedia lesson, cognitive load can be reduced by providing cues to the learner about how to select and organize the material—a technique called *signaling* (Lorch, 1989; Meyer, 1975). For example, Mautone and Mayer (2001) constructed a 4-min narrated animation explaining how airplanes achieve lift, which contained many extraneous facts and somewhat confusing graphics. Thus, the learner might engage in lots of incidental processing—by focusing on nonessential facts or nonessential aspects of the graphics. A signaled version guided the learner's cognitive processes of (a) selecting words by stressing key words in speech, (b) selecting images by adding red and blue arrows to the animation, (c) organizing words by adding an outline and headings, and (d) organizing images by adding a map showing which of three parts of the lesson was being presented. In the one study we conducted on signaling of a multimedia presentation (Mautone & Mayer, 2001, Experiment 3), students who received the signaled version of the narrated animation performed better on a subsequent test of problem-solving transfer than did students who received the unsignaled version. The effect size was .74. We refer to this result as a *signaling effect:* Students understand a multimedia presentation better when it contains signals concerning how to process the material. Although there is a substantial amount of research literature on signaling of text in printed passages (Lorch, 1989), Mautone and Mayer's study offers the first examination of signaling for narrated animations. Signaling seems to help in the process of selecting and organizing relevant information.

Type 4 Overload: Aligning and Eliminating Redundancy When the System is Overloaded by Incidental Processing Demands Attributable to How the Essential Material is Presented

Problem: One or both channels are overloaded by the combination of essential and incidental processing demands. The problem is the same in Type 3 and Type 4 overload—the learning task requires incidental processing—but the cause of the problem is different. In Type 3 over-

[5]Extraneous material may be related to the topic but does not directly support the educational goal of the presentation.

load the source of the incidental processing is that extraneous material is included in the presentation, but in Type 4 overload the source of the incidental processing is that the essential material is presented in a confusing way. For example, Type 4 overload occurs when on-screen text is placed at the bottom of the screen and the corresponding graphics are placed toward the top of the screen.

Solution: Aligning words and pictures. In Type 3 overload scenarios, incidental cognitive load was created by adding extraneous material. Another way to create incidental cognitive load is to misalign words and pictures on the screen, such as presenting an animation in one window with concurrent on-screen text in another window elsewhere on the screen. In this case—which we call a *separated presentation*—the learner must engage in a great deal of scanning to figure out which part of the animation corresponds with the words—creating what we call *incidental processing*. In eye-movement studies, Hegarty and Just (1989) showed that learners tend to read a portion of text and then look at the corresponding portion of the graphic. When the words are far from the corresponding portion of the graphic, the learner is required to use limited cognitive resources to visually scan the graphic in search of the corresponding part of the picture. The amount of incidental processing can be reduced by placing the text within the graphic, next to the elements it is describing. This form of presentation—which we call *integrated presentation*—allows the learner to devote more cognitive capacity to essential processing.

Consistent with this analysis, Moreno and Mayer (1999, Experiment 1) found that students who learned from integrated presentations (consisting of animation with integrated on-screen text) performed better on a problem-solving transfer test than did students who learned from separated presentations (consisting of animation with separated on-screen text). The effect size in this single study was .48. Similar effects have been found with text and illustrations in books (Mayer, 2001). We refer to this result as a *spatial contiguity effect:* Students understand a multimedia presentation better when printed words are placed near rather than far from corresponding portions of the animation. Thus, spatial alignment of words and pictures appears to be a valuable technique for reducing cognitive load. As you can see, aligning is similar to signaling in that it guides cognitive processing, eliminating the need for incidental processing. Aligning differs from signaling in that aligning applies to situations in which essential words and pictures are separated and signaling applies to situations in which extraneous material is placed within the multimedia presentation.

Solution: Eliminating redundancy. Another example of Type 4 overload occurs when a multimedia presentation consists of simultaneous animation, narration, and on-screen text. In this situation—which we call *redundant presentation*—the words are presented both as narration and simultaneously as on-screen text. However, the learner may devote cognitive capacity to processing the on-screen text and reconciling it with the narration—thus, priming incidental processing that reduces the capacity to engage in essential processing. In contrast, when the multimedia presentation consists of narrated animation—which we call *nonredundant presentation*—the learner is not primed to engage in incidental processing. In a series of three studies (Mayer et al., 2001, Experiments 1 and 2; Moreno & Mayer, 2002, Experiment 2) students who learned from nonredundant presentations performed better on problem-solving transfer tests than did students who learned from redundant presentations. The median effect size was .69, indicating that eliminating redundancy is a useful way to reduce cognitive load. We refer to this result as a *redundancy effect:* Students understand a multimedia presentation better when words are presented as narration rather than as narration and on-screen text. We use the term redundancy effect in a more restricted sense than Sweller (1999; Kalyuga, Ayres, Chandler, & Sweller, 2003). As you can see, eliminating redundancy is similar to weeding in that both involve cutting aspects of the multimedia presentation. They differ in that weeding involves cutting interesting but irrelevant material, whereas eliminating redundancy involves cutting an unneeded duplication of essential material.

When no animation is presented, students learn better from a presentation of concurrent narration and on-screen text (i.e., verbal redundancy) than from a narration-only presentation (Moreno & Mayer, 2002, Experiments 1 and 3). An explanation for this effect is that adding on-screen text does not overload the visual channel because it does not have to compete with the animation.

Type 5 Overload: Synchronizing and Individualizing When the System is Overloaded by the Need to Hold Information in Working Memory

Problem: One or both channels are overloaded by the combination of essential processing and representational holding. In the foregoing two sections, cognitive overload occurred when the learner attempted to engage in essential and incidental processing, and the solution was to reduce incidental processing through weeding and signaling (when extraneous material was included), or through aligning words and pictures or reducing redundancy (when the same essential material was presented in printed and spoken formats). In the fifth and final overload scenario, cognitive overload occurs when the learner attempts to engage in both essential processing (i.e., selecting, organizing, and integrating material that explains how the system works) and representational holding (i.e., holding visual and/or verbal representations in working memory during the learning episode).

For example, consider a situation in which a learner clicks on the *lightning* entry in a multimedia encyclopedia. First, a short narration is presented describing the steps in lightning formation; next, a short animation is presented depicting the steps in lightning formation. According to a cognitive theory of multimedia learning, this successive presentation can increase cognitive load because the learner must hold the verbal representation in working memory while the corresponding animation is being presented. In this situation, cognitive capacity must be used to hold a representation in working memory, thus depleting the learner's capacity for engaging in the cognitive processes of selecting, organizing, and integrating.

Solution: Synchronizing. A straightforward solution to the problem is to synchronize the presentation of corresponding visual and auditory material. When presentation of corresponding visual and auditory material is simultaneous, there is no need to hold one representation in working memory until the other is presented. This situation minimizes cognitive load. In contrast, when the presentation of corresponding visual and auditory material is successive, there is a need to hold one representation in one channel's working memory until the corresponding material is presented in the other channel. The additional cognitive capacity used to hold the representation in working memory can contribute to cognitive overload.

For example, in a series of eight studies carried out in our laboratory at UCSB (Mayer & Anderson, 1991, Experiments 1 and 2a; Mayer & Anderson, 1992, Experiments 1 and 2; Mayer, Moreno, Boire, & Vagge, 1999, Experiments 1 and 2; Mayer & Sims, 1994, Experiments 1 and 2), students performed better on tests of problem-solving transfer when they learned from simultaneous presentations (i.e., presenting corresponding animation and narration at the same time) than from successive presentations (i.e., presenting the complete animation before or after the complete narration). The median effect size was 1.30, indicating robust evidence for synchronizing as a technique for reducing cognitive load. We refer to this result as a *temporal contiguity effect*: Students understand a multimedia presentation better when animation and narration are presented simultaneously rather than successively.

Note that the temporal contiguity effect is eliminated when the successive presentation is broken down into bite-size segments that alternate between a few seconds of narration and a few seconds of corresponding animation (Mayer et al., 1999, Experiments 1 and 2; Moreno & Mayer, 2002, Experiment 2). In this situation, working memory is not likely to become overloaded because only a small amount of material is subject to representational holding.

Solution: Individualizing. When synchronization may not be possible, an alternative technique for reducing cognitive load is to be sure that the learners possess skill in holding mental representations in memory.[6] For example, high-spatial ability involves the ability to hold and manipulate mental images with a minimum of mental effort. Low-spatial learners may not be able to take advantage of simultaneous presentation because they must devote so much cognitive processing to hold mental images. In contrast, high-spatial learners are more likely to benefit from simultaneous presentation by being able to carry out the essential cognitive processes required for meaningful learning. Consistent with this prediction, Mayer and Sims (1994, Experiments 1 and 2) found that high-spatial learners performed much better on problem-solving transfer tests from simultaneous presentation than from successive presentation, whereas low-spatial learners performed at the same low level for both. Across two experiments involving a narrated animation on how the human respiratory system works, the median effect size was 1.13. We refer to this interaction as the *spatial ability effect*, and we note that individualization—matching high-quality multimedia design with high-spatial learners—may be a useful technique for reducing cognitive load.

CONCLUSION

Meeting the Challenge of Designing Instruction That Reduces Cognitive Load

A major challenge for instructional designers is that meaningful learning can require a heavy amount of essential cognitive processing, but the cognitive resources of the learner's information processing system are severely limited. Therefore, multimedia instruction should be designed in ways that minimize any unnecessary cognitive load. In this article we summarized nine ways to reduce cognitive load, with each load-reduction method keyed to an overload scenario.

Our research program—conducted at UCSB over the last 12 years—convinces us that effective instructional design depends on sensitivity to cognitive load which, in turn, depends on an understanding of how the human mind works. In this article, we shared the fruits of 12 years of programmatic research at UCSB and related research, aimed at contributing to cognitive theory (i.e., understanding the nature of multimedia learning) and building an empirical database (i.e., research-based principles of multimedia design).

Theory. We began with a cognitive theory of multimedia learning based on three core principles from cognitive science, which we labeled as dual channel, limited capacity, and active processing (shown in Table 1). Based on the cognitive theory of multimedia learning (shown in Figure 1), we de-

[6] Individualization is not technically a design method for reducing cognitive load but rather a way to select individual learners who are capable of benefitting from a particular multimedia presentation.

rived predictions concerning various methods for reducing cognitive load. In conducting dozens of controlled experiments to test these predictions, we were able to refine the theory and offer substantial empirical support. Thus, the seemingly practical search for load-reducing methods of multimedia instruction has contributed to theoretical advances in cognitive science—a well-supported theory of how people learn from words and pictures. Overall, our approach has been based on the idea that the best way to improve instruction is to begin with a research-based understanding of how people learn.

Database. Our search for theory-based principles of instructional design led us to conduct dozens of well-controlled experiments—thereby producing a substantial research base (summarized in Table 3). For each of our recommendations for how to reduce cognitive load, we see the need to conduct multiple experiments. In some cases when we report only a single preliminary study (i.e., segmenting, signaling, and aligning) more empirical research is needed. Clear and replicated effects are the building blocks of both theory and practice. Overall, our approach has been based on the idea that the best way to understand how people learn is to test theory-based predictions in the context of student learning scenarios.

Future directions. Additional research is needed on the measurement of cognitive load (cf. Brüncken, Plass, & Leutner, 2003; Paas, Tuovinen, Tabbers, & Van Gerven, 2003). In particular, we need ways to gauge (a) cognitive load experienced by learners, (b) the cognitive demands of instructional materials, and (c) the cognitive resources available to individual learners. Although we hypothesize that our nine recommendations reduce cognitive load, it would be useful to have direct measures of cognitive load.

In our research, concise narrated animation fostered meaningful learning without creating cognitive overload. However, additional research is needed to examine situations in which certain kinds of animation can overload the learner (Schnotz, Boeckheler, & Grzondziel, 1999) and to determine the role of individual differences in visual and verbal learning styles in influencing cognitive overload (Plass, Chun, Mayer, & Leutner, 1998; Riding, 2001). In addition, it would be worthwhile to examine whether the principles of multimedia learning apply to the design of online courses that require many hours of participation, to problem-based simulation games, and to multimedia instruction that includes on-screen pedagogical agents (Clark & Mayer, 2003).

In short, our program of research convinces us that the search for load-reducing methods of instruction contributes to cognitive theory and educational practice. Research on multimedia learning promises to continue to be an exciting venue for educational psychology.

ACKNOWLEDGMENT

This research was supported by Grant N00014–01–1–1039 from the Office of Naval Research.

REFERENCES

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

Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38,* 53–61.

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

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

Clark, R. C., & Mayer, R. E. (2003). *E-learning and the science of instruction.* San Francisco: Jossey-Bass.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48,* 163–189.

Hegarty, M., & Just, M. A. (1989). Understanding machines from text and diagrams. In H. Mandl & J. R. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 171–194). Amsterdam: Elsevier.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38,* 23–31.

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

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

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

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

Mayer, R. E. (2002). *The promise of educational psychology: Vol. 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., & 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., 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., & Wetzell, K. (2002). Fostering understanding of multimedia messages through pre-training: Evidence for a two-stage theory of mental model construction. *Journal of Experimental Psychology: Applied, 8,* 147–154.

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. 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,* 389–460.

Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In D. Berliner & 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.
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). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology, 92,* 117–125.
Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94,* 156–163.
Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based multimedia learning: 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.
Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38,* 63–71..
Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.
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.
Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction, 12,* 61–86.
Riding, R. (2001). The nature and effects of cognitive style. In R. J. Sternberg & L. Zhang (Eds.), *Perspectives on thinking, learning, and cognitive styles* (pp. 47–72). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Schnotz, W., Boeckheler, J., & Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. *European Journal of Psychology of Education, 14,* 245–265.
Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.
van Merriënboer, 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.