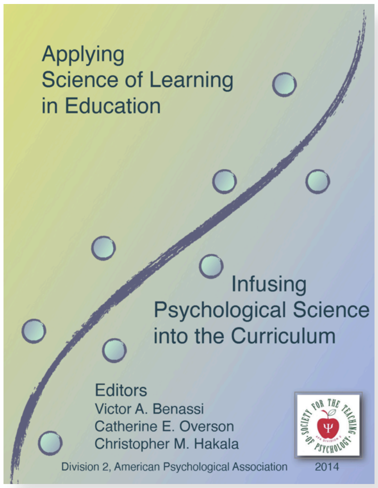


Research-Based Principles for Designing Multimedia Instruction

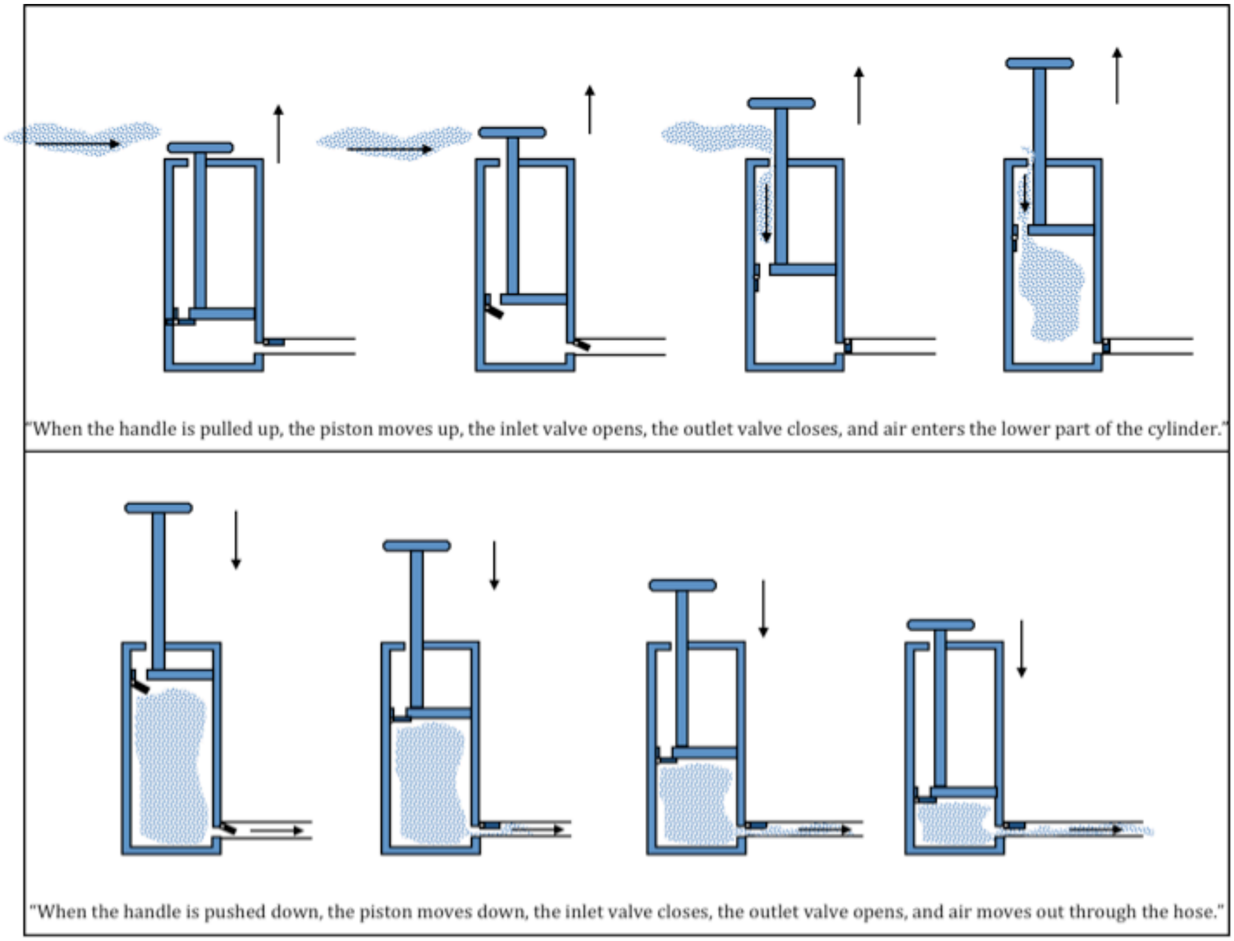
Richard E. Mayer
University of California, Santa Barbara

Overview of Multimedia Instruction

People learn more deeply from words and graphics than from words alone. This assertion can be called the *multimedia principle*, and it forms the basis for using *multimedia instruction*—that is, instruction containing words (such as spoken text or printed text) and graphics (such as illustrations, charts, photos, animation, or video) that is intended to foster learning (Mayer, 2009).

For example, Figure 1 shows frames from a narrated animation on how a tire pump works. In this case the words are spoken and the graphics are presented as an animation. Other examples include textbook lessons presented on paper, slideshow presentations presented face-to-face, captioned video presented via computer, or educational games and simulations presented on hand-held devices. Regardless of presentation medium, what makes all these examples of multimedia instruction is that they use words and graphics to promote learning.

Figure 1: Frames from a Narrated Animation on How a Tire Pump Works



What is the evidence for the multimedia principle? In a series of 11 experimental comparisons my colleagues and I have found that students perform much better on a transfer test when they learn from words and graphics than from words alone (e.g., narration and animation versus narration alone, or text and illustrations versus text alone), yielding a median effect size of d = 1.39 (Mayer, 2009). For example, students are better able to answer troubleshooting questions about tire pumps after viewing the narrated animation shown in Figure 1 than after solely hearing the narration (Mayer & Anderson, 1991). In short, research shows that multimedia instruction has the potential to greatly improve student understanding of how things work, including tire pumps, brakes, generators, and lightning (Butcher, in press; Mayer, 2009).

Although instruction has traditionally emphasized verbal modes of presentation (such as lectures and printed books), recent advances in graphics technology now allow more widespread incorporation of visual modes of presentation including illustrations, charts, photos, animation, and video in presentations and in interactive venues such as games and simulations. However, not all graphics are equally effective so careful research is needed to pinpoint principles of multimedia instructional design. The goal of this chapter is to provide a brief overview of 12 research-based principles for how to design effective instruction that uses words and graphics.

An important starting point is to examine principles that are based on an understanding of how people learn from words and graphics. Figure 2 summarizes the cognitive theory of multimedia learning (Mayer, 2009, in press-a), which is based on three core assumptions based on the science of learning (Mayer, 2011): *dual channel assumption*—people have separate channels for processing visual and verbal material (Paivio, 1986); *limited capacity assumption*—people can process only a limited amount of material in a channel at any one time (Baddeley, 1999); and *active processing assumption*—meaningful learning occurs when learners select relevant material, organize it into a coherent structure, and integrate it with relevant prior knowledge (Wittrock, 1989; Mayer, 2009).

Figure 2 represents memory stores as rectangles: *sensory memory,* which temporarily holds incoming images and sounds; *working memory*, which allows for mentally manipulating a small amount of the incoming visual and verbal material; and *long-term memory,* which is the learner's permanent storehouse of knowledge. Figure 2 represents cognitive processing as arrows: *selecting*, which transfers some of the incoming images and sounds to working memory for additional processing; *organizing,* which organizes the images into a pictorial model and the words into a verbal model in working memory; and *integrating,* which connects the models with each other and with relevant knowledge activated from long-term memory. A multimedia message enters the cognitive system through the learner's ears and eyes. The top row represents the verbal channel (into which spoken words and sounds enter) and the bottom row represents the visual channel (into which graphics and printed words enter), although in working memory printed words can be converted into sounds and images can be converted into spoken words.

Figure 2: Cognitive Theory of Multimedia Learning

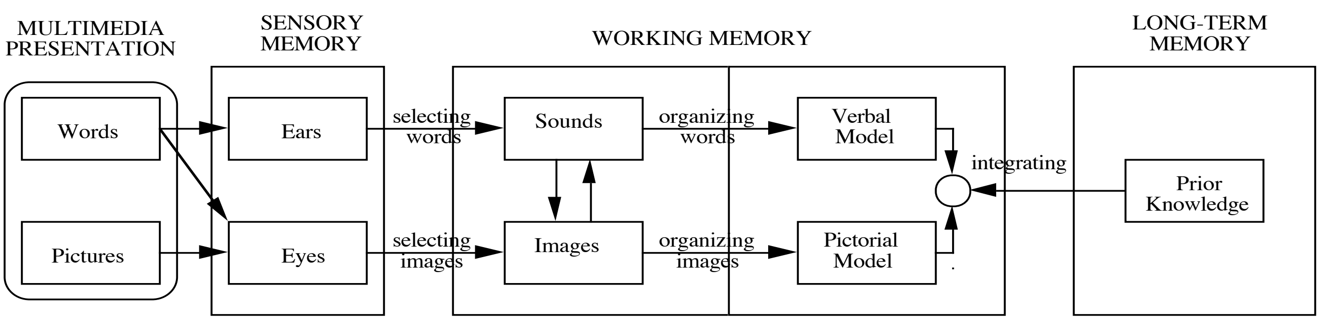


Table 1 summarizes three kinds of processing that can occur during multimedia instruction—*extraneous processing,* which drains limited cognitive processing capacity without contributing to learning; *essential processing,* which involves selecting relevant information and organizing it as presented in working memory; and *generative processing*, which involves making sense of the material by reorganizing it into a coherent structure and integrating it with relevant prior knowledge. This analysis is similar to that proposed in cognitive load theory (Sweller, Ayres, & Kalyuga, 2011) and suggests the need for three kinds of instructional design goals—*reducing extraneous processing*, when extraneous processing and required essential processing exceed the learner's cognitive capacity; *managing essential processing*, when the required essential processing exceeds the learner's cognitive capacity; and *fostering generative processing*, when the learner has processing capacity available but chooses not to exert the effort to use it for making sense of the material. These three types of goals form the basis for three kinds of instructional design principles for multimedia learning, which are presented in the next section.

Table 1: Three Kinds of Cognitive Processing During Learning

Cognitive processing	Description	Instructional goal
Extraneous	Not related to instructional goal, caused by poor instructional design	Reduce extraneous processing
Essential	Aimed at representing essential material, caused by complexity of material	Manage essential processing
Generative	Aimed at making sense of essential material, caused by learner's effort	Foster generative processing

Research on Design Principles for Multimedia Instruction

In this section I examine 12 research-based principles for how to design multimedia, including five principles aimed at reducing extraneous processing, three principles aimed at managing essential processing, and four principles aimed at fostering generative processing. For each principle, we conducted a meta-analysis (Ellis, 2010) and computed the median effect size based on Cohen's d (Cohen, 1988; see Bertsch & Pesta, in this volume, for a definition of the meaning of d and a brief description of meta-analysis). Following Hattie (2009), we consider any effect size greater than d = 0.40 to be educationally important.

Reduce Extraneous Processing

Table 2 summarizes five principles aimed at reducing extraneous processing: the coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. The table specifies the number of experimental tests in which positive results were obtained and provides the median effect size based on a meta-analysis by Mayer and Fiorella (in press). A learner experiences *extraneous overload* when essential cognitive processing required to understand the essential material in a multimedia message and extraneous cognitive processing required to process extraneous material exceeds the learner's cognitive capacity. These five principles are intended to address the problem of extraneous overload.

Table 2: Five Research-Based Principles Based on Reducing Extraneous Processing

Principle	Description	Tests	Effect size
Coherence	Delete extraneous material	22 of 23	0.86
Signaling	Highlight essential material	25 of 29	0.41
Redundancy	Don't add onscreen captions to narrated graphics	16 of 16	0.86
Spatial contiguity	Place printed words near corresponding part of graphic	22 of 22	1.10
Temporal contiguity	Present spoken words at same time as corresponding graphics	9 of 9	1.22

The Coherence Principle

People learn more deeply from a multimedia message when extraneous material is excluded rather than included. The rationale for the coherence principle is that people are better able to focus on the essential material if we eliminate extraneous material that could distract them. This principle was supported in 22 out of 23 experimental tests, yielding a median effect size of 0.86. For example, students who learned from a multimedia lesson on how a virus causes a cold performed better on a transfer test if the lesson did not contain seductive details—sentences that gave interesting but irrelevant facts about viruses (Mayer, Griffith, Jurkowitz, & Rothman, 2008). Similarly, students who learned from a narrated animation on lightning formation performed better on a transfer test if the lesson did not also contain short video clips depicting lightning strikes (Mayer, Heiser, & Lonn, 2001). Concerning boundary conditions, reviews of the coherence principle suggest the effects may be strongest for learners with low rather than high working memory capacity, when the lesson is system-paced rather than learner paced, and when the extraneous material is highly interesting rather than neutral (Mayer & Fiorella, in press; Rey, 2012).

The Signaling Principle

People learn more deeply from a multimedia message when cues are added that highlight the organization of the essential material. The rationale for the signaling principle is that people will learn more efficiently if the lesson is designed to call their attention to the important material in the lesson and how it is organized. This principle was supported in 25 out of 29 experimental tests, yielding a median effect size of 0.41. Signaling of the verbal material includes using an outline, headings, highlighting (such as underlining) and pointer words (such as first, second, third). Signaling of visual material includes arrows, flashing, and spotlighting. For example, in a narrated animation on how an airplane achieves lift, students performed better on a transfer test if the narration included an initial outline, headings, and voice emphasis on key words (Mautone & Mayer, 2001). Concerning boundary conditions, Mayer and Fiorella (in press) report that the signaling principle may apply most strongly when the learner might otherwise be overwhelmed with extraneous processing—such as, for low-knowledge learners rather than high-knowledge learners, for complex material rather than simple material, and when it used sparingly rather than excessively.

The Redundancy Principle

People learn more deeply from graphics and narration than from graphics, narration, and on-screen text. The rationale is that with redundant presentations people may waste precious processing capacity by trying to reconcile the two verbal streams of information or may focus on the printed words rather than the relevant portions of the graphics. This principle was supported in 16 out of 16 experimental tests, yielding a median effect size of 0.86. For example, Moreno and Mayer (2002) reported that students performed better on a transfer test about lightning if they received a narrated animation about lightning formation rather than the same narrated animation with concurrent onscreen text inserted at the bottom of the screen. Concerning boundary conditions, the redundancy effect can be diminished or even reversed when the learners are experienced, the on-screen text is short, or the material lacks graphics (Mayer, 2009; Mayer & Fiorella, in press).

The Spatial Contiguity Principle

People learn more deeply from a multimedia message when corresponding printed words and graphics are presented near rather than far from each other on the page or screen. The rationale is that spatial contiguity helps learners build connections between corresponding words and graphics. This principle was supported in 22 out of 22 experimental tests, yielding a median effect size of 1.10. For example, Moreno and Mayer (1999) found that students performed better on a transfer test after viewing an animation about lightning in which printed words were placed next to the part of the lightning system they described than when printed words were placed at the bottom of the screen as a caption. Similar results are reported in a meta-analysis by Ginns (2006). Mayer and Fiorella (in press) report there is preliminary evidence that the spatial contiguity principle may be strongest for low prior knowledge learners, non-redundant text and pictures, complex lessons, and interactive formats.

The Temporal Contiguity Principle

People learn more deeply from a multimedia message when corresponding graphics and narration are presented simultaneously rather than successively. The rationale is that temporal contiguity helps learners build connections between corresponding words and graphics. This principle was supported in 9 out of 9 experimental tests, yielding a median effect size of 1.22. For example, students performed better on a transfer test when they received a narrated animation on how a tire pump works than when they heard the narration before or after the animation (Mayer and Anderson, 1991). Researchers have noted that the temporal contiguity principle may apply most strongly for high rather than low spatial ability learners, when lessons are long rather than short, and when the lesson is system paced rather than learner paced (Mayer, 2009; Mayer & Fiorella, in press).

Manage Essential Processing

Table 3 summarizes three principles aimed at managing essential processing: the segmenting, pre-training, and modality principles. The table specifies the number of experimental tests in which positive results were obtained and provides the median effect size based on a meta-analysis by Mayer and Pilegard (in press). These principles are intended to address the instructional problem of *essential overload,* which can occur when a fast-paced multimedia lesson contains material that is complicated for the learner. A learner experiences *essential overload* when the amount of essential cognitive required to understand the multimedia instructional message exceeds the learner's cognitive capacity.

Table 3: Three Research-Based Principles Based on Managing Essential Processing

Principle	Description	Tests	Effect size
Segmenting	Break lesson into learner-paced parts	10 of 10	0.79
Pre-training	Present characteristics of key concepts before lesson	13 of 16	0.75
Modality	Use spoken words rather than printed words	52 of 61	0.76

The Segmenting Principle

People learn more deeply when a multimedia message is presented in learner-paced segments rather than as a continuous unit. The rationale is that segmenting allows people to fully process one step in the process before having to move onto the next one. This principle was supported in 10 out of 10 experimental tests, yielding a median effect size of 0.79. For example, Mayer and Chandler (2001) found that students performed better on a transfer test if a narrated animation about lightning was broken into 16 10-second segments in which students could press a "continue" button to go on to the next segment. The review revealed that some potential boundary conditions are that segmenting may have stronger effects for learners with low rather than high working memory capacity and for low achieving rather than high achieving learners (Mayer & Pilegard, in press).

The Pre-training Principle

People learn more deeply from a multimedia message when they have learned the names and characteristics of the main concepts. The rationale is that pre-training allows students to focus on the causal connections in the multimedia explanation because they already know the names and characteristics of the key elements. This principle was supported in 13 out of 16 experimental tests, yielding a median effect size of 0.75. For example, students performed better on a transfer test based on a narrated animation on how brakes work if before the lesson they were introduced to the names and characteristics of key components mentioned in the lesson such as the piston and brake shoe (Mayer, Mathias, & Wetzell, 2002). However, an important boundary condition is that the pre-training principle may not apply to high prior knowledge learners, perhaps because they are less likely to experience essential overload (Mayer, 2009; Mayer & Pilegard, in press).

The Modality Principle

People learn more deeply from a multimedia message when the words are spoken rather than printed. The rationale is that the modality principle allows learners to off-load some of the processing in the visual channel (i.e., the printed captions) onto the verbal channel, thereby freeing more capacity in the visual channel for processing the animation. This principle was supported in 53 out of 61 experimental tests, yielding a median effect size of 0.76. For example, Moreno and Mayer (1999) found that students performed better on a transfer test after receiving a narrated animation on lightning formation than after receiving the same animation with on-screen captions that contained the same words as the narration. Similar results are reported in a meta-analysis by Ginns (2005). As the most studied principle in the list, research shows that the modality principle should not be taken to mean that spoken words are better than printed words in all situations. Some important boundary conditions reported in some studies are that printed word may be effective when the verbal material contains technical terms, is in the learner's second language, or is presented in segments that are too large to be held in the learner's working memory (Mayer, 2009; Mayer & Pilegard, in press).

Foster Generative Processing

Table 4 summarizes four principles aimed at fostering generative processing: the personalization, voice, embodiment, and image principles. The table specifies the number of experimental tests in which positive results were obtained and provides the median effect size based on a meta-analysis by Mayer (in press-b). These principles are intended to use social cues to prime the learner's motivation to exert effort to make sense of the material. Social cues in a multimedia message such as conversational style, voice, and gesture may prime a sense of social presence in learners that leads to deeper cognitive processing during learning and hence better test performance.

Table 4: Four Research-Based Principles Based on Fostering Generative Processing

Principle	Description	Tests	Effect size
Personalization	Put words in conversational style rather than formal style	14 of 17	0.79
Voice	Put words in human voice rather than machine voice	4 of 5	0.69
Embodiment	Have onscreen agent use human-like gestures and movements	11 of 11	0.36
Image	Do not necessarily put static image of agent on the screen	9 of 14	0.20

The Personalization Principle

People learn more deeply when the words in a multimedia presentation are in conversational style rather than formal style. The rationale for this technique is that conversational style can prime a sense of social presence in the learner, which causes the learner to try harder to make sense of what the instructor is saying by engaging in appropriate cognitive processing during learning, leading to learning outcomes that are better able to support problem-solving transfer. This principle was supported in 14 out of 17 experimental tests, yielding a median effect size of $d = 0.79$. For example, Mayer, Fennell, Farmer and Campbell (2004) found that students performed better on a transfer test after receiving a narrated animation on how the human respiratory system works when conversational wording was used (e.g., "your lungs," or "your nose") rather than formal style (e.g., "the lungs" or "the nose"). Similar results are reported in a meta-analysis by Ginns, Martin, and Marsh (in press). Some important boundary conditions are that the personalization principle may not apply for high-achieving students or long lessons.

The Voice Principle

People learn more deeply when the words in a multimedia message are spoken in a human voice rather than in a machine voice. Human voice is intended to prime a sense of social presence in learners. For example, Mayer, Sobko, and Mautone (2003) found that students performed better on a transfer test after receiving a narrated animation lightning that used a human voice rather than a machine synthesized voice. This principle was supported in 4 out of 5 experimental comparisons, with a median effect size of $d = 0.69$. Research shows that a possible boundary condition is that the voice principle may not apply when there are negative social cues such as low embodiment.

The Embodiment Principle

People learn more deeply when onscreen agents display human-like gesturing, movement, eye contact, and facial expression. Human-like action is intended to create a sense of social presence with the instructor. In 11 out of 11 experimental comparisons, people performed better on transfer tests when they learned from a high-embodied agent than from a low-embodied agent, yielding a median effect size of *d* = 0.36. For example, Mayer and DaPra (2012) found that students performed better on a transfer test after viewing an online slideshow that was narrated by an onscreen agent who used human-like gesture, facial expression, eye gaze, and movement than an onscreen agent who did not move, gesture, gaze, or show expression. A possible boundary condition is that the embodiment principle may not apply when there are negative social cues such as machine voice.

The Image Principle

People do not necessarily learn more deeply from a multimedia presentation when the speaker's image is on the screen rather than not on the screen. Having a static image may cause distraction that detracts from any social benefits. For example, Mayer, Dow, and Mayer (2003) found that adding the image of an onscreen character did not improve learning much (*d* = 0.19) from a narrated animation on how electric motors work. This principle is based on 14 experimental tests in which half produced negative or negligible effects, yielding a median effect size of *d* = 0.20.

Practical Application of Design Principles for Multimedia Instruction

The principles summarized in this chapter are based on research and grounded in cognitive theory, but more work is needed to better delineate the boundary conditions under which the principles apply. In particular, most of the supporting research involves short-term laboratory studies, so it is useful to determine the degree to which the principles apply in more authentic learning situations such as in schools or work training. For example, a promising step in this direction involves a recent finding by Issa et al. (2013) showing that redesigning a medical school classroom slideshow lecture on shock based on multimedia principles resulted in improvements on an immediate transfer test (*d* = 0.76) and a delayed transfer test (*d* = 1.17).

The principles presented in this chapter are intended to apply to a range of instructional scenarios ranging from textbooks to face-to-face slide show presentations to computer-based lessons to interactive games and simulations. Within a classroom, these principles apply to the design of classroom printed materials, computer-based exercises and simulations, and face-to-pace instruction including slideshow presentations.

For example, suppose you wished to apply research-based multimedia design principles to improve a short slideshow on how a solar cell works for presentation in an environmental science class. First, in deference to the coherence principle, you might decide to prune interesting but irrelevant material that you had downloaded from the Internet, including photos of a solar cell installation in a desert in southern California and a short video clip you found in which Al Gore envisions the coming environmental disaster. In short, you work to weed out images and words that are not essential to explaining how a solar cell works, including eliminating entire slides or parts of slides.

In deference to the signaling principle, you place a heading at the top of the remaining 10 slides, which succinctly describe the step in the process being depicted. In addition, you use arrows to show the movement of electrons within the solar cell, corresponding to the description in the text.

In line with the segmenting principle, you break the description in each slide into a few very short descriptions of actions rather than one long paragraph.

In line with the spatial contiguity principle, you place the text describing the action next to the corresponding part of the graphic (such as putting “Electrons move across the barrier” next to an arrow from one side of the barrier to another) rather than at the bottom of the graphic.

Based on the pre-training principle you begin with a slide that that depicts the key elements (such as the “positive side” and “negative side” of the solar cell) along with a verbal label next to each key element, perhaps connected with a line.

Corresponding to the redundancy principle, you do not simply read the text on the slides, but rather elaborate on it, using complete sentences.

In this case, the strictest interpretation modality principle and the redundancy principle can be modified by including a minimal number of words on the screen—mainly, to help highlight the main points and to concretize technical or unfamiliar terms.

In line with the personalization principle, you use first and second person style (such as saying, “Let’s see how a solar cell works.” rather than “This lesson tells how a solar cell works.”).

Consistent with the embodiment principle, you stand next to the slide as you talk, pointing out what you are talking about and maintaining a smile and eye contact with the audience.

Based on temporal contiguity, you are careful to choreograph your speech so that it corresponds to what you are pointing to in the graphic.

In line with the voice principle, you practice to make sure you use a pleasant, flowing voice that exudes confidence and warmth.

In line with the image principle, you remove a distracting logo from each slide which shows Al Gore’s face looking down along with the slogan: “LIVE GREEN.”

All in all, this example demonstrates how to accomplish your goal of removing material that primes extraneous processing, helping learners understand a complex system by using techniques such as segmenting and pre-training, and fostering deeper processing by creating a feeling of social conversation. I will consider this chapter to be a success to the extent that instructors and instructional designers are able to improve student learning by adapting the principles presented in this chapter and other evidence-based principles (see Mayer, in press-c).

Acknowledgement

Preparation of this chapter was supported by a grant from the Office of Naval Research.

References

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

Butcher, K. (in press). The multimedia principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning.* New York: Cambridge University Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Mahwah, NJ: Erlbaum.

Ellis, P. D. (2010). *The essential guide to effect sizes.* New York: Cambridge University Press.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 15*, 313-332.

Ginns, P. (2006). Integrating information: A meta-analysis of spatial contiguity and temporal contiguity effects. *Learning and Instruction, 16,* 511-525.

Ginns, P., Martin, A.J., & Marsh, H.M. (in press). Designing instructional text in a conversational style: A meta-analysis. *Educational Psychology Review.*

Hattie, J. (2009). *Visible learning.* New York: Routledge.

Issa, N., Mayer, R. E., Schuller, M., Wang, E., Shapiro, M. B., & DaRosa, D. A. (2013). Teaching for understanding in medical classrooms using multimedia design principles. *Medical Education, 47,* 388-396.

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. (2009). *Multimedia learning* (2nd ed). New York: Cambridge University Press.

Mayer, R. E. (2011). *Applying the science of learning*. Upper Saddle River, NJ: Pearson.

Mayer, R. E. (in press-a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning.* New York: Cambridge University Press.

Mayer, R. E. (in press-b). Principles based on social cues: Personalization, voice, image, and embodiment principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning.* New York: Cambridge University Press.

Mayer, R. E. (Ed.). (in press-c). *The Cambridge handbook of 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., & 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., & DaPra, C. S. (2012). An embodiment effect in computer-based learning with animated pedagogical agent. *Journal of Experimental Psychology: Applied, 18,* 239-252.

Mayer, R. E., Dow, G., & Mayer, R. E. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology, 95,* 806-813.

Mayer, R. E., Fennell, S., Farmer, L., & Campbell, J. (2004). A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. *Journal of Educational Psychology, 96*, 389-395.

Mayer, R. E. , & Fiorella, L. (in press). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Mayer, R. E., Griffith, E., Jurkowitz, I. T. N., & Rothman, D. (2008). Increased interestingness of extraneous details in multimedia science presentation leads to decreased learning. *Journal of Experimental Psychology: Applied, 14,* 329-339.

Mayer, R. E., Heiser, H., & 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. , & Pilegard, C. (in press). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology, 95*, 419-425.

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. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology, 94,* 598-610.

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

Rey, G. D. (2012). A review and a meta-analysis of the seductive detail effect. *Educational Research Review, 7,* 216-237.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory.* New York: Springer

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