Evidence-Based Principles for How to Design Effective Instructional Videos

Richard E. Mayer

University of California, Santa Barbara (USA)

Target article for JARMAC

Date submitted: March 3, 2021

Revision submitted: March 24, 2021

Word count: 4173

Contact: Richard E. Mayer

905 Crown Avenue

Santa Barbara, CA 93111

mayer@ucsb.edu

805-451-2290

Abstract

Drawing on research conducted by Mayer and colleagues (Mayer, 2020), this article examines evidence-based principles for how to design effective instructional videos and shows how they are grounded in cognitive theories of learning and instruction. Principles include multimedia (present words and graphics), coherence (avoid extraneous material in slides and script), signaling (highlight key material), redundancy (do not add captions that repeat the spoken words), spatial contiguity (place printed text next to corresponding part of graphic), temporal contiguity (present corresponding visual and verbal material at the same time), segmenting (break a complex slide into progressively presented parts), pre-training (provide pre-training in the names and characteristics of key concepts), modality (present words as spoken text). personalization (use conversational language), voice (use appealing human voice), image (do not display static image of instructor's face), embodiment (display gesturing instructor), and generative activity (add prompts for generative learning activity).

Keywords: multimedia learning, video lectures, instructional video, animation, computer-based learning

General Audience Summary

Suppose you wanted to create a video of a lecture for students to access online. In the instructional video, you are standing next to a board that displays a series of slides. How can we design the video lecture to help students learn? This article describes a research-based theory of how people learn from multimedia presentations, such as instructional video, and then summarizes 14 evidence-based principles for how to design effective instructional video. For example, research shows that students learn better when extraneous comments and unneeded visual material are excluded (i.e., coherence principle) and key material is highlighted (i.e., signaling principle), when a complex slide is broken down into a progression of added parts (i.e., segmenting principle), and when the instructor uses conversational language (i.e., personalization principle) and gesture when speaking (i.e., embodiment principle).

The Case for Instructional Video

The COVID-19 pandemic has caused massive changes in the ways that education is being delivered around the globe. For example, instead of face-to-face class meetings in which teachers and students interact in person, we now are relying on instructional video as a primary instructional medium. In this article, I focus on instructional videos in which instructors present slides (or animations) as they lecture. Instructional video is a type of *multimedia instructional message*--that is, a presentation consisting of words and visuals that is intended to promote learning (Mayer, 2020; Mayer, Fiorella, & Stull, 2020). In the case of instructional video, the words are presented as the lecturer's narration or as onscreen printed text and the visuals are in the slides or animation shown on the screen.

For example, consider the situation summarized in Figure 1 in which an instructor stands next to a screen that shows a progression of slides as she explains the concept of binomial probability (Lawson et al., in press). What can be done to improve the instructional effectiveness of this 10-minute long video lecture? In this article, I examine evidence-based principles for how to design effective instructional video, derived from our research program on multimedia instructional design (Mayer, 2020). For the past 30 years, my colleagues and I at the University of California, Santa Barbara have been examining how to design effective multimedia instruction--including instructional video--based on research evidence and grounding in a cognitive theory of learning and instruction. Our aim is to contribute to the broader search around the globe for principles of multimedia instructional design (Mayer & Fiorella, in press-a). In this article I share the fruits our research efforts as an example of a sustained effort to apply the science of learning to education. First, I summarize a research-based theory of how people learn and how to help people learn from multimedia lessons such as instructional video. I summarize evidence-based principles for how to design instructional video, including five principles for how to guide learners away from distraction and towards the essential material during learning, three principles for how to help students manage processing of essential material that is complex for them, and five principles for how to motivate students to exert effort to make sense of the essential material.

Grounding in a Cognitive Theory of Multimedia Learning

Table 1 lists three principles from the science of learning that are relevant to a cognitive theory of multimedia learning: the dual channels, limited capacity, and active learning principles (Mayer, 2020). The dual channels principle is that learners have separate information processing channels for visual and verbal material (Paivio, 1986, 2006). The limited capacity principle is that learners can processes only a few pieces of information in each channel at any one time (Baddeley, 1999; Sweller, Ayres, & Kalyuga, 2011). The active processing principle is that learners must engage in appropriate cognitive during learning, including attending to the relevant incoming information (i.e., selecting), mentally organizing it into a coherent representation (i.e., organizing), and relating it to relevant knowledge activated from long-term memory (i.e., integrating; Fiorella & Mayer, 2015; Wittrock, 1989). The challenge for instructional designers is to effectively use the learner's visual and verbal channels to allow for active cognitive processing while not overwhelming the learner's processing capacity.

These principles about how the human mind works help to shape a basic framework for describing how people learn from multimedia instructional messages such as video lectures, which I call the cognitive theory of multimedia learning (Mayer, 2020, in press-a). Figure 2 summarizes the cognitive theory of multimedia learning, which contains dual channels (represented as the top and bottom rows), has limited capacity (represented as the rectangle containing working memory), and depends on active cognitive processing (represented as arrows for selecting, organizing, and integrating). First, the learner receives a multimedia instructional message, which enters sensory memory through the learner's eyes and ears. The sensory input in sensory memory decays rapidly, but can be transferred to working memory by paying attention (represented by the *selecting* arrows). In working memory, which has limited capacity, the learner can mentally arrange the visual material into a pictorial model and the verbal information into a verbal model (represented by the *organizing* arrows). Finally, the learner can integrate the pictorial and verbal models with each other and with relevant prior knowledge activated from long-term memory (represented by the *integrating* arrows). The resulting new knowledge can then be stored in long-term memory, which has unlimited capacity for meaningful material. As can be seen, the arrows represent the cognitive processes needed for meaningful learning--selecting, organizing, and integrating; whereas the boxes represent the memory stores--sensory memory, working memory, and long-term memory.

Grounding in a Cognitive Theory of Multimedia Instruction

Table 2 lists three demands on the learner's cognitive resources during instruction (including during viewing a video lecture): extraneous processing, essential processing, and generative progressing (Mayer, 2020, in press-a). Extraneous processing refers to cognitive processing during learning that does not serve the instructional objective of the lesson; it can be caused by poor instructional design, such as including extraneous graphics and text in the lesson. Essential processing is cognitive processing during learning aimed at building a mental representation of the essential presented material in working memory; the level of essential processing is caused by the complexity of the material for the learner. Generative processing is cognitive processing during learning aimed at making sense of the material; the level of generative processing depends on the learner's motivation to exert effort to understand the material. The learner's cognitive capacity for building knowledge in working memory is limited, so when cognitive capacity is used for one of these types of processing it cannot be used for another.

Figure 3 summarizes three instructional scenarios that require different types of instructional design solutions: extraneous overload, essential overload, and generative underuse (Mayer, 2020, in press-a). In extraneous overload, learners use so much of their cognitive capacity on extraneous processing that they do not have enough remaining capacity to engage in the essential and generative processing that is needed to build an appropriate mental representation of the material. This can happen when aspects of the lesson are particularly distracting, such as when a video lesson has a lot of highly interesting but irrelevant images. In essential overload, the lesson is so complicated for the learner that processing it overwhelms the learner's cognitive capacity. This can happen when a video lecture is full of information presented at a fast pace to introductory-level learners. In generative underuse, learners have cognitive capacity available, but do not use it to engage in generative processing during learning. This can occur when learners are viewing an instructional video that they find uninteresting and of little value to them.

Based on this analysis, I have derived three goals for instructional design of multimedia lessons: reducing extraneous processing, managing essential processing, and fostering generative processing. Table 3 summarizes five instructional design principles for reducing extraneous processing, three principles for managing essential processing, and five principles for fostering generative processing. In the following sections, for each of these principles I provide a description and example along with empirical evidence and theoretical justification. My focus is on design principles that improve learning outcomes, particularly transfer performance--that is, the ability to apply the presented material to solve new problems. For example, to assess the learner's understanding of a how a pump works, we can ask the learner to write answers for a troubleshooting question (e.g., what could be wrong if a pump does not work) or a redesign question (e.g., how could we make a pump more reliable or powerful).

The Multimedia Principle

I begin with the most central and overarching principle for the design of instructional video--the *multimedia principle* (Mayer, 2020, in press-b). Consider a brief video lecture in which an instructor sits at her computer in front of the camera and speaks the following explanation of how a bicycle tire pump works: "When the handle is pulled up, the piston moves up, the inlet valve opens, the outlet valve closes, and air enters the lower part of the cylinder. When the handle is pushed down, the piston moves down, the inlet valve closes, the outlet valve opens, and air moves out through the hose." In this case, the video lecture consists of a talking head producing a narration. When we test students, they do not perform well on answering transfer questions, such as, "Suppose you pull up and push down on the handle several times, but no air comes out. What could be wrong?" or "What could you do to make a much more effective?" Next, in order to boost learning, suppose you add graphics to your narration, perhaps in the form of a short narrated animation as depicted in Figure 4. In this case, students perform much better on a transfer test. Overall, in 3 out of 3 experimental comparisons, we found that students performed better on a transfer posttest after receiving a narrated animation than after receiving narration alone, yielding a median effect size of d = 1.90 (Mayer & Anderson, 1991, 1992). These kinds of findings support the *multimedia principle*: People learn better from words and graphics than from words alone (Mayer, in press-a). This principle is highlighted at the top of Table 3. Consistent with the cognitive theory of multimedia learning, multimedia presentations are more likely to allow learners to build both verbal and pictorial representations that can be integrated, which is a core cognitive process required for meaningful learning.

Although adding graphics to words has been shown to boost learning outcomes, not all multimedia messages (including instructional videos) are equally effective. In the following sections, I explore evidence-based principles for how to design effective instructional videos.

Principles for Reducing Extraneous Processing: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles

Based on the cognitive theory of multimedia learning, an important instructional design goal is to reduce extraneous processing (Fiorella & Mayer, in press-a). The first major section of Table 3 lists five ways to reduce extraneous processing, which I call the coherence, signaling, redundancy, spatial contiguity, and temporal continuity principles.

First, as shown in Table 3, one way to reduce extraneous processing from an instructional video is to weed out non-essential words and graphics, which I call the *coherence principle.* Consider a narrated animation on how lightning storms develop, as partially depicted in Figure 5. To spice up the lesson, we can add interesting but irrelevant facts about lightning, such as information about how many people are struck by lightning each year, and we can intersperse interesting but irrelevant video clips such as an ambulance rushing to pick up a golfer who was struck by lightning on a golf course. When we add this interesting but irrelevant material, people score lower on a transfer test (Mayer, Heiser, & Lonn, 2001). Overall, across 3 out of 3 experiments involving slideshows or animation, people performed better when extraneous material was excluded rather than included, yielding a median effect size of d = 0.80 (Mayer, Heiser, & Lonn, 2001; Mayer, Griffith, Naftaly, & Rothman, 2008). Similarly, across two experiments, removing background music from a narrated animation on lightning or car braking systems also improved transfer test performance, yielding a median effect size of d = 0.95 (Moreno & Mayer, 2000a).

The next technique for reducing extraneous processing in Table 3 is to highlight the essential information in an instructional video, which I call the *signaling principle*. Consider a narrated animation explaining how an airplane achieves lift. To emphasize the key terms, the speaker can stress them through speech, which is a form of verbal signaling. This technique improved transfer test performance in two experiments, with a median effect size of d = 0.65 (Mautone & Mayer, 2001). Next consider a video showing an onscreen agent standing next to a graphic as she lectures about neural transmission. To guide the learner's visual attention toward the relevant part of a graphic, we can use red coloring and/or pointing, which is a form of visual signaling. This is exemplified in Figure 6. In a collection of 10 experiments, these techniques tended to increase transfer test performance, with a median effect size of d = 0.73 (Li, Wang, Mayer, & Liu, 2019; Wang, Li, Mayer, & Liu, 2018; Xie, Wang, Mayer, & Zhou, 2019).

Another technique for reducing extraneous processing in Table 3 is to present graphics and narration (as exemplified in the left side of Figure 7) rather than graphics, narration, and identical on-screen text (as exemplified in the right side of Figure 7), which we call the *redundancy principle*. When we remove redundant onscreen text from the lightning lesson, this improves transfer test performance across four experiments, yielding a median effect size of d = 0.80 (Mayer, Heiser, & Lonn, 2001; Mayer & Johnson, 2008; Moreno & Mayer, 2002). According to the cognitive theory of multimedia learning, onscreen text can create extraneous processing when learners try to look back and forth between the printed caption and the animation.

The next way to reduce extraneous processing is to place printed words next to the part of the graphic they refer to rather than as a caption away from the graphic, which I call the *spatial contiguity principle*. For example, suppose a student is viewing annotated animation on lightning formation with the words presented as a caption at the bottom of the screen (as shown in the right side of Figure 8) or as words integrated next to the corresponding part of the graphic (as shown in the left side of Figure 8). In two experiments, students performed better on transfer tests about lightning when printed words were placed next to the corresponding part of the graphic rather than separated as a caption, yielding a median effect size of $d = 0.65$ (Makransky, Terkildsen, & Mayer, 2019; Moreno & Mayer, 1999). Spatial contiguity reduces the need for the learner to scan back and forth between the caption and the graphic, thereby reducing extraneous processing.

The final technique for reducing extraneous processing listed in Table 3 is to synchronize the narration with the video or animation so that material on the screen corresponds to what the instructor is saying. I call this the *temporal continuity principle*. In 8 out of 8 experiments, students performed better on a transfer test if they received narration in sync with corresponding animation that if they received narration before or after animation for lessons on pumps, brakes, lightning, and lungs, yielding a median effect size of $d = 1.31$ (Mayer & Anderson, 1991, 1992; Mayer & Sims, 1994; Mayer, Moreno, Boire, & Vagge, 1999). When the presentation is successive rather than simultaneous, learners need to hold verbal information in working memory until corresponding visual information is presented (or vice versa), which wastes limited working memory capacity.

Principles for Managing Essential Processing: Segmenting, Pre-training, and Modality Principles

Even if we design instructional video based on principles for reducing extraneous processing, the essential content of the lesson may be complex for the learner to the point of overwhelming the learner's limited processing capacity. In this case, we need instructional design techniques for managing essential processing (Mayer & Fiorella, in press-b). The next section of Table 3 lists three techniques for managing the processing of essential material--the segmenting, pre-training, and modality principles.

First, as shown in Table 3, one way to manage essential processing is to break a complex presentation into manageable segments whose pace is controlled by the learner. For example, consider a narrated slideshow on how geographic information systems (GIS) works. When presented one slide at a time, the amount of information on the screen can be overwhelming. To reduce the load on the learner's information processing, we could allow the learner to click an arrow key to progressively add to the slide as the instructor speaks. This type of segmenting improved transfer test performance across two experiments (Mayer, Howarth, Kaplan & Hanna, 2018; Mayer, Wells, Parong, & Howarth, 2019), and overall across 7 out of 7 experiments, various forms of segmenting of slideshows resulted in better transfer test performance as compared to a continuous presentation, yielding a median effect size of d = 0.67 (Mautone & Mayer, 2007; Mayer & Chandler, 2001; Mayer, Dow, & Mayer, 2003; Sung & Mayer, 2013). Segmenting allows the learner to build a mental representation of one part of the material before moving on to the next.

The second way to manage essential processing listed in Table 3 is to provide pre-training in the names and characteristics of the key elements in the lesson. For example, consider a narrated animation on how a car's braking system works. The lesson is presented at a fast pace and contains terms that may be unfamiliar to the learner, such as *piston, master* *cylinder, brake shoe, brake drum,* and so on. In order to reduce the load on the learner's working memory during the lesson, we could provide a pre-training exercise in which the learner can click on each component in a diagram of the braking system and see the name of the part and a description of what it does. For example, when the learner clicks on the piston, the piston moves forward and back in a spotlighted area and text says: "This is the piston in the master cylinder. It can either move forward or back." Across two experiments, students who received this pre-training before the brakes lesson performed better on a transfer than those who did not receive pre-training, with a median effect size of $d = 0.85$ (Mayer, Mathias, & Wetzell, 2002).

The third technique for managing essential processing is to present words in an animation or video in spoken form rather than as onscreen text, which I call the *modality principle.* Across, 7 of 7 lessons involving learning in one's first language, students performed better on a transfer test from a multimedia lesson involving spoken text rather than printed text, with a median effect size of $d = 1.02$ (Harskamp, Mayer, Suhre, & Jansma, 2007; Mayer, Wells, Parong, & Howarth, 2019; Mayer & Moreno, 1998; Moreno & Mayer, 1999). However, in learning in one's second language, students learned better with printed words rather than spoken words from a video lesson on Antarctica, reflecting an important boundary condition for the modality principle (Lee & Mayer, 2018). The transient nature of spoken text makes it most appropriate when the words are familiar to the learner.

Principles for Fostering Generative Processing: Personalization, Voice, Image, and Embodiment Principles

Even if we can design instruction that minimizes extraneous processing and manages essential processing so learners have processing capacity available, they may not use that capacity to make sense of the material, perhaps because they are not motivated to learn (Fiorella

& Mayer, in press-b, in press-c). In this section, I explore techniques for priming learners to use their available cognitive capacity to engage in generative processing during learning based on the personalization, voice, image, embodiment, and generative activity principles.

As you can see in the bottom portion of Table 3, the first way to foster generative processing is to use conversational language rather than formal language, which I call the *personalization principle*. For example, in 5 out of 5 experiments, students performed better on a transfer test if the instructor used conversational language rather than formal language in a narrated animation on how the human respiratory system works (e.g., "your mouth" rather than "the mouth") or how lightning storms develop (e.g., "your choice" rather than "the choice"), yielding a median effect size of $d = 1.00$ (Mayer, Fennell, Farmer & Campbell, 2004; Moreno & Mayer, 2000b). Personalized language is intended to help the learner feel that the instructor is working with them, which can prime stronger motivation to exert effort to understand what the instructor is saying.

A similar idea is that people learn better when instructor speaks with a pleasant human voice rather than a machine synthesized voice, which I call the *voice principle*. This principle was upheld in 6 out of 7 experiments, involving a narrated slideshow on how solar cells work (Mayer & DaPra, 2012), how lightning storms develop (Mayer, Sobko, & Mautone, 2003), and how to solve word problems (Atkinson, Mayer, & Merrill, 2005), yielding a median effect size of $d = 0.74$.

Next, consider the role of having the instructor's static image on the screen during a narrated slideshow. In two experiments, including a static image of the instructor standing motionless next to a progression of slides resulted in poorer transfer test performance than not having any image in the screen, with a median effect size of d = -0.25 (Mayer & DaPra, 2012;

Wang, Li, Mayer, & Liu, 2018). This supports the image principle, which states that people do not learn better from narrated graphics when a static image of the instructor is added to the screen. A motionless image may seem unnatural to learners and distract them from the content in the lesson.

Although a static instructor may not be helpful, let's explore what happens when we add an embodied instructor--that is, an instructor who draws as she lectures, maintains eye contact as she lectures, uses appropriate gesture during lecturing, or is filmed from a first-person perspective. These features constitute what I call the *embodiment principle*. For example, consider an instructional video in which the instructor draws graphics on a board as she lectures about the Doppler effect (as in the left side of Figure 9) or simply points towards already drawn graphics (as in the right side of Figure 9). Fiorella and Mayer (2016) found that low-knowledge students performed on a transfer test if they saw an instructor drawing as she lectures, with an effect size of $d = 0.58$.

As another example of the embodiment principle, consider an instructor who faces the class and maintains eye contact as she lectures about human kidneys using a transparent board (as shown in the left side of Figure 10) or faces away and does not maintain eye contact (as shown in the right side of Figure 10). Fiorella, Stull, Kuhlmann, and Mayer (2020) found that students performed better on a transfer test when the instructor maintained eye contact than when she did not, and similar results were obtained with instructional video lessons on chemistry (Stull, Fiorella, & Mayer, 2018; Stull, Fiorella, Gainer, & Mayer, 2018), yielding a median effect size of $d = 0.60$ based on four experiments.

As yet another example of the embodiment principle, across eight experiments, students performed better on a transfer test if they had viewed a narrated slideshow in which the onscreen agent engaged in human-like gesture as she talked rather than if she stood still in lessons involving solar cells (Mayer & DaPra, 2012) and neural transmission (Wang, Li, Mayer, & Liu, 2019), yielding a median effect size of $d = 0.76$. As a final example, consider a video demonstration of how to build an electric circuit on a circuit board that is filmed from a first-person perspective (as shown in the left panel of Figure 11) or from a third-person perspective (as shown in the right panel of Figure 11). Across two experiments, we found that students performed better on a transfer test if they have viewed an instructional video demonstration filmed from a first-person perspective, with a median effect size of $d = 0.53$ (Fiorella, van Gog, Hoogerheide, & Mayer, 2017). Thus, there is converging evidence that instructor embodiment is an important consideration in creating instructional video.

Finally, the last technique for promoting fostering processing listed in Table 3 is to include prompts to engage in generative learning activities, that is, activities that the learner engages in during instruction that are intended to promote learning (Fiorella & Mayer, 2015, in press-c). For example, suppose a student views an online series of animated slides that explains how global warming works. For some students, we pause the lesson after each slide and ask the learner to type in a one-sentence explanation of what was just presented. This technique, which can be called learning-by-explaining (Fiorella & Mayer, 2015), has been shown to improve transfer test performance in 3 of 3 experiments, with a median effect size of d = 0.68 (Lawson & Mayer, in press). Learning by explaining can be effective when it causes students to reflect on what they are learning.

The Future of Instructional Video

In this brief review, I examined 14 evidence-based principles for how to design effective instructional videos (including narrated slideshows and animations), and showed how they are grounded in cognitive theories of learning and instruction. Future research is needed to establish a broader research base, to determine the boundary conditions under which the principles most strongly apply, to pinpoint the cognitive and motivational processes that mediate the effects, and to examine the degree to which the principles apply in authentic learning environments. This effort is an example of applying the science of learning to education and training (Clark & Mayer, 2016; Mayer, 2011).

Acknowledgement: Preparation of this article was supported by Grant N0014612046 from the Office of Naval Research. This article is based on Mayer (2020).

References

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology*, *30*, 117-139.

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

Clark, R. C., & Mayer, R. E. (2016). *e-Learning and the science of instruction* (4th ed). San Francisco: Wiley.

Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity.* New York: Cambridge University Press.

Fiorella, L., & Mayer, R. E. (2016). Effects of observing the instructor draw diagrams on learning from multimedia messages. *Journal of Educational Psychology, 108*, 528-546.

Fiorella, L., & Mayer, R. E. (in press-a). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed). New York: Cambridge University Press.

Fiorella, L., & Mayer, R. E. (in press-b). Principles based on social cues in multimedia learning: Personalization, voice, embodiment, and image principles. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed). New York: Cambridge University Press.

Fiorella, L., & Mayer, R. E. (in press-c). The generative activity principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed). New York: Cambridge University Press.

Fiorella, L., Stull, A., Kuhlmann, S. L., & Mayer, R. E. (2020). Fostering generative learning from video lessons: Benefits of instructor-generated drawings and learner-generated explanations. *Journal of Educational Psychology*, *112*, 895-906.

Fiorella, L., van Gog, T., Hoogerheide, V., & Mayer, R. E. (2017). It's all a matter of perspective: Viewing first-person video modeling examples promotes learning of an assembly task. *Journal of Educational Psychology, 109*, 653-665.lawson

Harskamp, E., Mayer, R. E., Suhre, C., & Jansma, J. (2007). Does the modality principle for multimedia learning apply to science classrooms? *Learning and Instruction, 17*, 465-477.

Lawson, A. P., & Mayer, R. E. (in press). Benefits of writing an explanation during pauses in multimedia lessons. *Educational Psychology Review, 00,* 000-000.

Lawson, A. P., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (in press). Do learners recognize and relate to the emotions displayed by virtual instructors? *International Journal of Artificial Intelligence in Education, 00,* 000-000.

Lee, H., & Mayer, R. E. (2018). Fostering learning from instructional video in a second language. *Applied Cognitive Psychology, 32,* 648-654.

Li, W., Wang, F., Mayer, R.E., & Liu, H. (2019). Getting the point: Which kinds of gestures by pedagogical agents improve multimedia learning? *Journal of Educational Psychology, 111,* 1382-1395.

Makransky, G., Terkildsen, T., & Mayer, R. E. (2019). Role of subjective and objective measures of cognitive processing during learning in explaining the spatial contiguity effect. *Learning and Instruction, 61,* 23-34.

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

Mautone, P. D, & Mayer, R. E. (2007). Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology, 99*, 640-652.

Mayer, R. E. (2011). *Applying the science of learning.* Boston, MA: Pearson.

Mayer, R. E. (2020). *Multimedia learning* (3rd ed). New York: Cambridge University Press.

Mayer, R. E. (in press-a). Cognitive theory of multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed). New York: Cambridge University Press.

Mayer, R. E. (in press-b). The multimedia principle. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed). 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., & 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 agents. *Journal of Experimental Psychology: Applied, 18,* 239-252.

Mayer, R. E., Dow, G., & Mayer, S. (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-a). *The Cambridge handbook of multimedia learning* (3rd ed). New York: Cambridge University Press.

Mayer, R. E., & Fiorella, L. (in press-b). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed). New York: Cambridge University Press.

Mayer, R. E., Fiorella, L., & Stull, A. (2020). Five ways to increase the effectiveness of instructional video. *Educational Technology Research and Development, 68,* 837-852.

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

Mayer, R. E., Howarth, J., Kaplan, M., & Hanna, S. (2018). Applying the segmenting principle to online geography lessons. *Educational Technology Research and Development, 66,* 563-577.

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

Mayer, R. E., & Johnson, C. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology, 100*, 380-386.

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. *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, 86,* 389-401.

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.

Mayer, R. E., Wells, A., Parong, J., & Howarth, J. (2019). Learner control of the pacing of an online slideshow: Does segmenting help? *Applied Cognitive Psychology, 33*, 930-935.

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. (2000a). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia messages. *Journal of Educational Psychology, 92,* 117-125.

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

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94,* 156-163.

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

Paivio, A. (2006). *Mind and its evolution: A dual coding approach.* Mahwah, NJ: Erlbaum.

Stull, A., Fiorella, L., Gainer, M., & Mayer, R. E. (2018). Using transparent whiteboards to boost learning from online STEM lectures. *Computers & Education, 20,* 146-159.

Stull, A., Fiorella, L., & Mayer, R. E. (2018). An eye-tracking analysis of instructor presence in video lectures. *Computers in Human Behavior,* 88, 263-272.

Sung, E., & Mayer, R. E. (2013). Online multimedia learning with mobile devices and desktop computers: An experimental test of Clark's methods-not-media hypothesis. *Computers in Human Behavior, 29*, 639-647.

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

Wang, F., Li, W., Mayer, R. E., & Liu, H. (2018). Animated pedagogical agents as aids in multimedia learning: Effects on eye-fixations during learning and learning outcomes. *Journal of Educational Psychology, 110*, 250-268.

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

Xie, H., Wang, F., Mayer, R. E., & Zhou, Z. (2019). Coordinating visual and auditory cueing in multimedia learning. *Journal of Educational Psychology, 111,* 235-255.

Table 1

Three Principles from the Science of Learning

Principle	Description
Dual channels	People have separate channels for processing visual and verbal information.
Limited capacity	People can process only a few items in each channel in working memory at any one time.
Active learning	Meaning learning requires that learners engage in appropriate cognitive processing during learning, including selecting, organizing, and integrating.

Table 2

Three Demands on the Learner's Cognitive Capacity During Learning

Name	Description	Cause
Extraneous processing	Processing that does not serve the instructional objective	Poor instructional design
Essential processing	Processing aimed at mentally representing the essential presented material	Complexity of the material
Generative processing	Processing aimed at making sense of the material	Motivation to exert effort

Table 3

Evidence-Based Principles for the Design of Instructional Videos

Principle	Description	Example
Multimedia	Present words and graphics rather than words alone.	Video contains both narration and graphics.
REDUCING EXTRANEOUS PROCESSING		
Coherence	Avoid extraneous material in slides and script.	Slides do not contain huge logos or colorful backgrounds.
Signaling	Highlight key material. .	Lecturer points to elements in slide as she mentions them.
Redundancy	Do not add captions that repeat the narration.	Video does not contain subtitles (unless words are technical or in the learner's second language).
Spatial contiguity	Place printed text next to corresponding part of graphic.	Slides contain graphics with words placed next to corresponding parts.
Temporal contiguity	Present visual material at same time as corresponding narration.	The lecturer's narration is in sync with material on the slide.
MANAGE ESSENETIAL PROCESSING		
Segmenting	Break a complex lesson into progressive parts under the control of the learner.	Allow the learner to press a button to see the next segment of a slide and hear the accompanying narration.
Pre-training	Provide pre-training in the names and characteristics of key concept.	Video contains an introduction involving the names and characteristics of key concepts.
Modality	Present words as spoken text.	Video includes instructor's voice.
FOSTER GENERATIVE PROCESSING		
Personalization	Use conversational language.	Lecturer speaks in first and second person using "I", "we", "you", "us", and/or "let's."

Voice	Use appealing human voice.	Lecturer speaks with friendly human voice that displays positive emotion.
Image	Do not display static image of the instructor's face.	Video does not have a window with a photo of the instructor's face.
Embodiment	Display gesturing instructor.	Lecturer writes and draws on board as she lectures. Lecturer maintains eye contact as she lectures. Lecturer displays dynamic gestures as she lectures. Demonstration is filmed from first-person perspective.
Generative activity	Insert generative learning activities.	At pauses in the video the learner is asked to type in a brief explanation of the foregoing segment.

Figure 1

Screenshot from an instructional video lesson

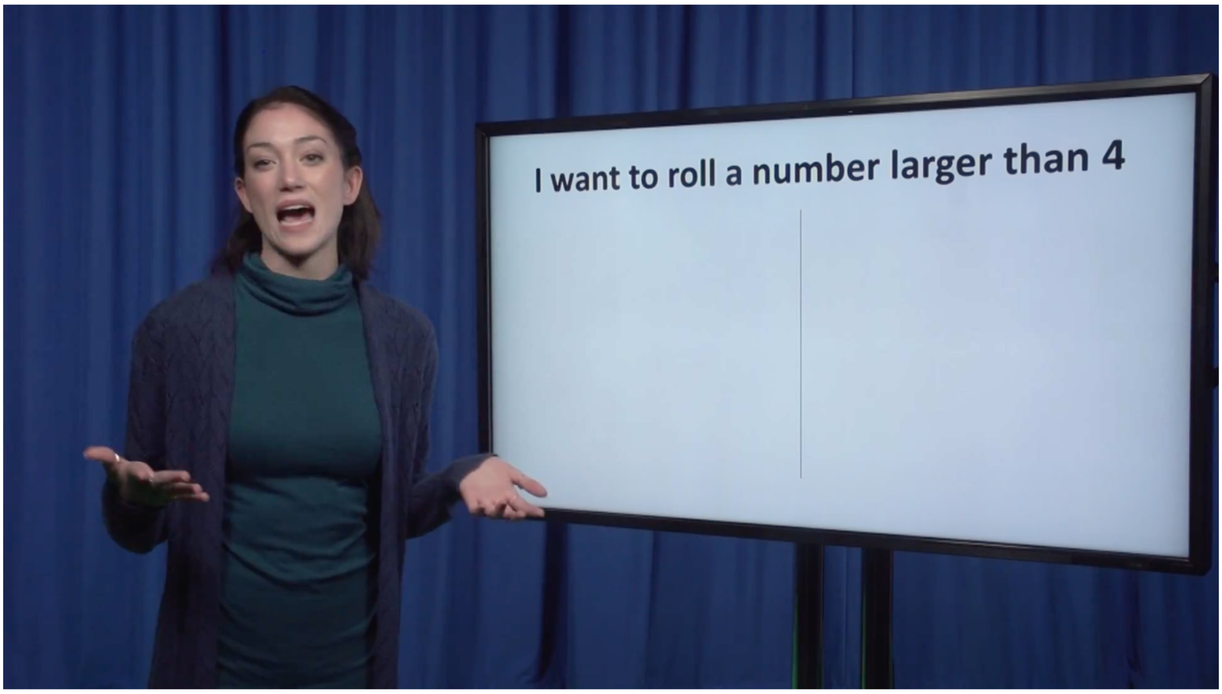


Figure 2

Cognitive theory of multimedia learning

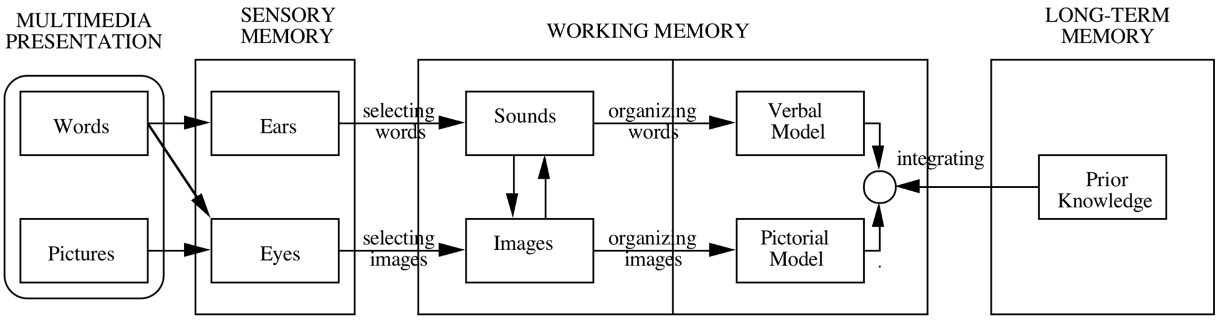


Figure 3

Three instructional scenarios

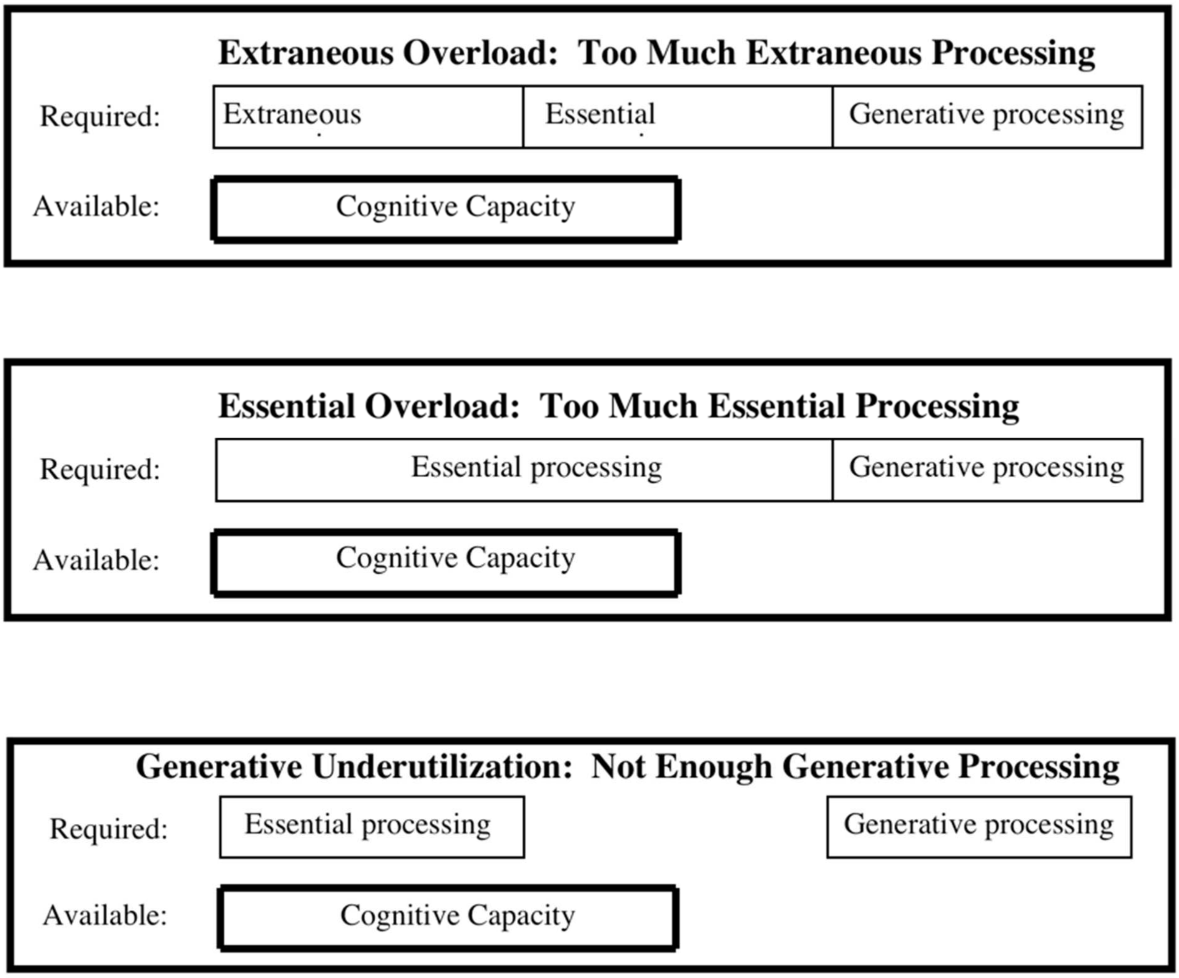


Figure 4. Frames from a narrated animation on how a bicycle tire pump works.

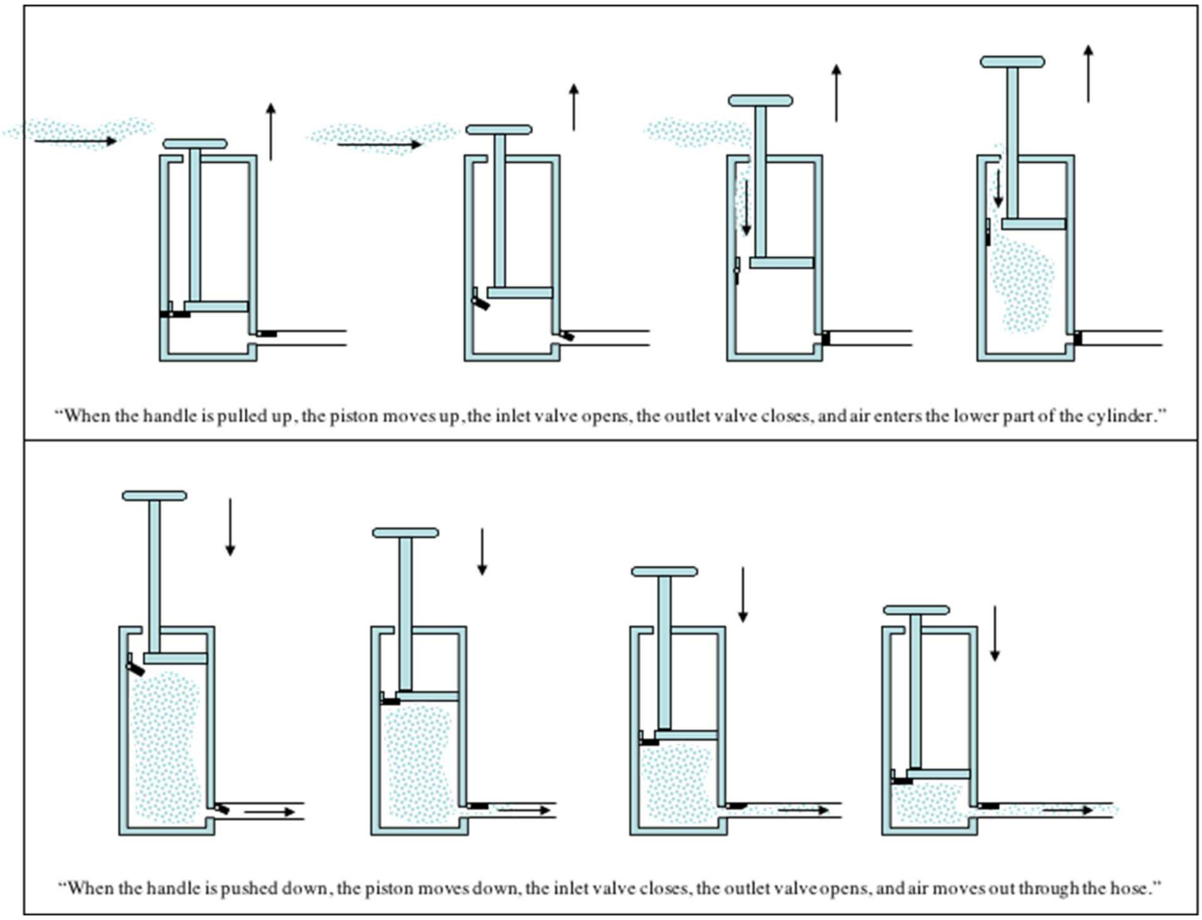


Figure 5. Selected frames from a narrated animation on lightning formation.

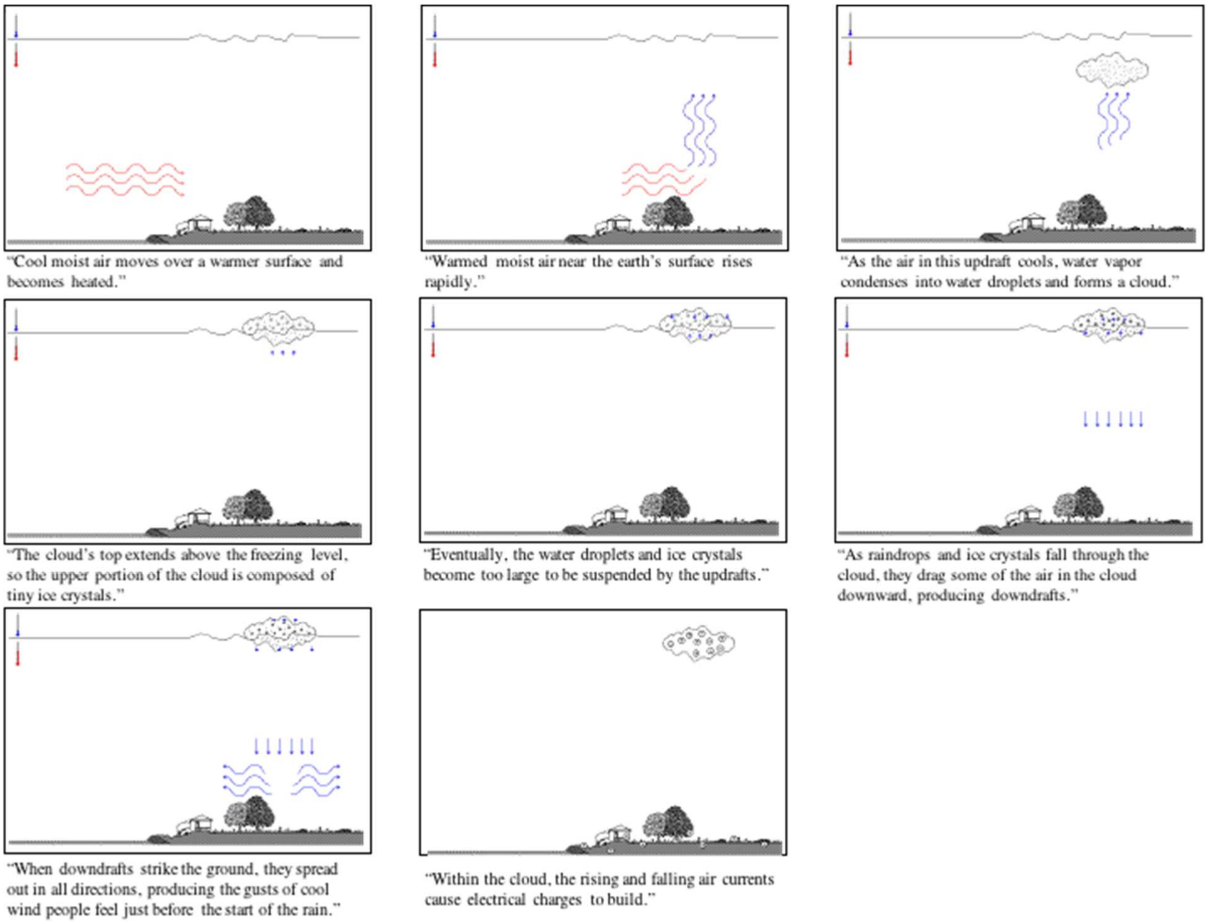


Figure 6. Screenshot from narrated animation on neural transmission with an agent's pointing gestures.

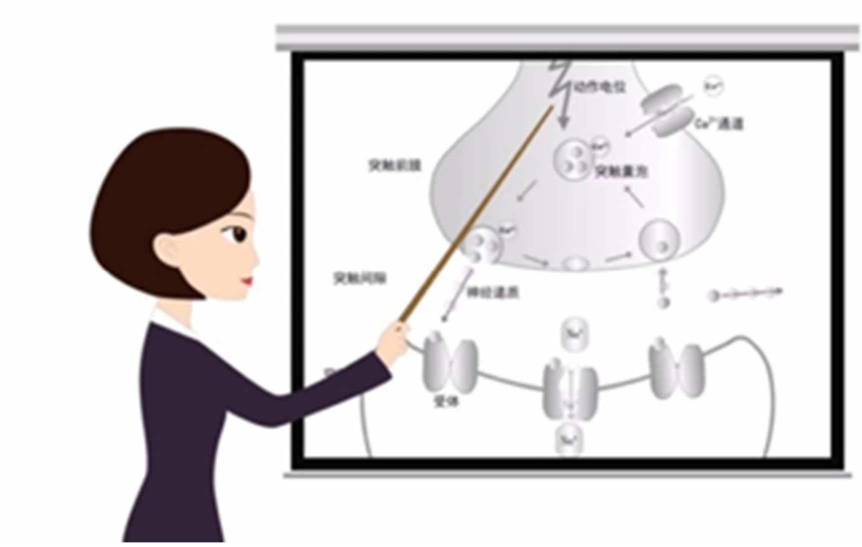


Figure 7. Frames from narrated animation and narrated animation with onscreen text.

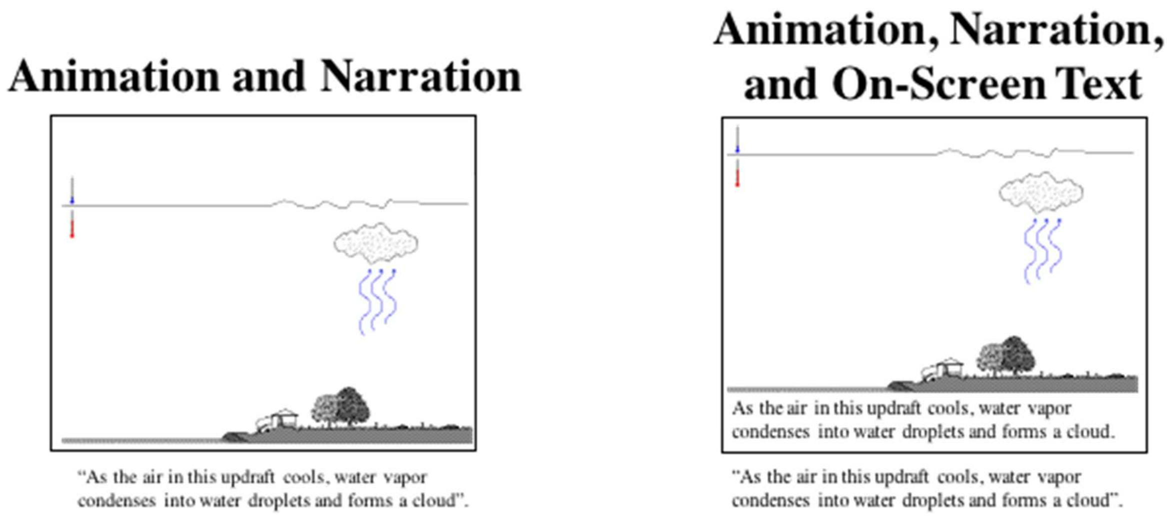


Figure 8. Frames from lightning lessons with words integrated or separated from graphics.

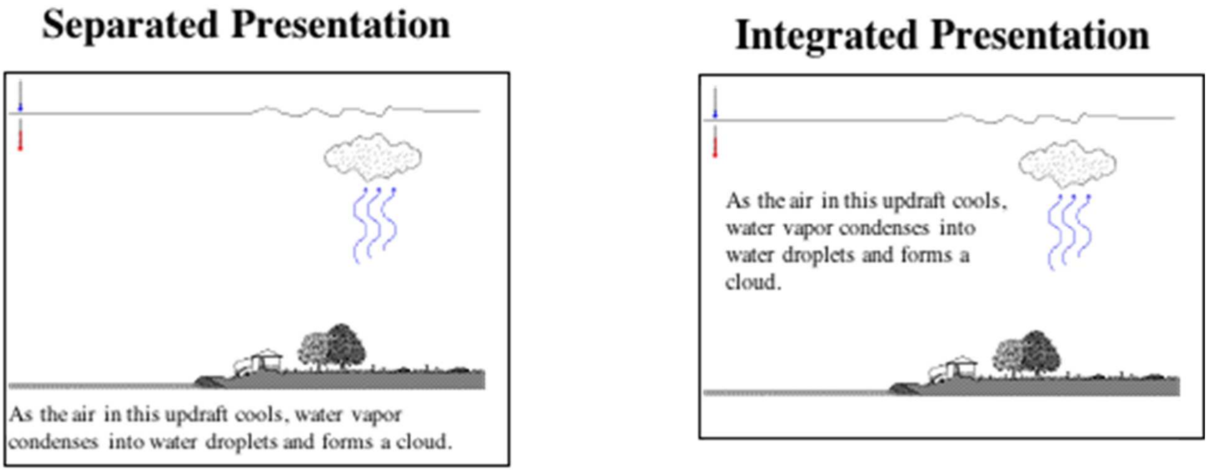


Figure 9. Frame from video on Doppler effect with instructor drawing as she lectures or points to already drawn illustrations.

Explain already drawn graphics: Draw graphics as you explain:

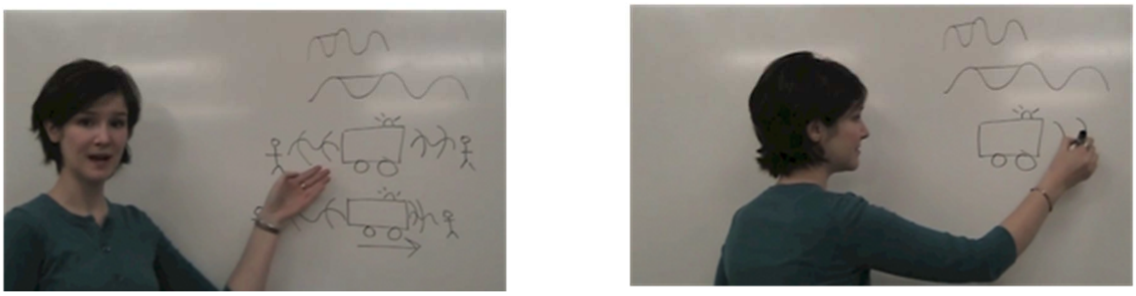
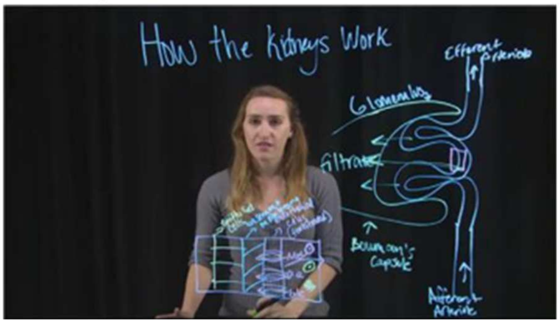


Figure 10. Frame from video on human kidneys with instructor who maintains eye contact or does not maintain eye contact.

Instructor faces the class with transparent whiteboard:



Instructor faces board with conventional whiteboard:

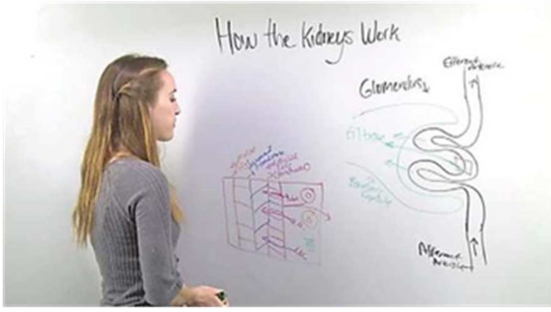


Figure 11. Frame from video on circuit building demonstration from a first- or third-person perspective.