Review Article

doi: 10.1111/jcal.12197

Journal of Computer Assisted Learning

Using multimedia for e-learning

R.E. Mayer

Department of Psychological and Brain Sciences, University of California, CA, USA

Abstract

This paper reviews 12 research-based principles for how to design computer-based multimedia instructional materials to promote academic learning, starting with the multimedia principle (yielding a median effect size of $d = 1.67$ based on five experimental comparisons), which holds that people learn better from computer-based instruction containing words and graphics rather than words alone. Principles aimed at reducing extraneous processing (i.e., cognitive processing that is unrelated to the instructional objective) include coherence ($d = 0.70$), signalling ($d = 0.46$), redundancy ($d = 0.87$), spatial contiguity ($d = 0.79$) and temporal contiguity ($d = 1.30$). Principles for managing essential processing (i.e., mentally representing the essential material) include segmenting ($d = 0.70$), pre-training ($d = 0.46$) and modality ($d = 0.72$). Principles for fostering generative processing (i.e., cognitive processing aimed at making sense of the material) include personalization ($d = 0.79$), voice ($d = 0.74$) and embodiment ($d = 0.36$). Some principles have boundary conditions, such as being stronger for low- rather than high-knowledge learners.

Keywords e-learning, instructional design, multimedia learning, science of learning.

Computer-assisted learning (CAL) dates back to programmed instruction delivered via mainframe computers in the 1950s and 1960s but has grown to include a new generation of CAL that can be called e-learning (Cuban, 1986; Saettler, 1990, 2004). e-Learning is defined as instruction that is delivered via a digital device that is intended to promote learning (Clark & Mayer, 2016). The digital device can be any electronic device that is controlled by a computer chip, including a desktop computer, laptop computer, tablet, smartphone, game console or wearable devices such as head mounted virtual reality displays.

e-Learning plays a significant and increasing role in education and training, due in part to making learning available nearly anytime and anywhere. In the domain of worker training, the percentage of worker training available via e-learning has increased from 11 to 39% from 2002 to 2013 while the percentage of training available in instructor-led classrooms has decreased from 70 to 55% (ATD, 2014). In addition to increased use in traditional K–12 and college venues, e-learning has found its way into many creative forms of alternative and informal learning, ranging from websites such as PBS Kids or ABC Mouse aimed at helping children learn academic content outside of school, online self-help services such as Khan Academy or Lynda.com aimed at training specific skills, online universities such as University of Phoenix or Kaplan University, and MOOCS such as Udacity, Coursera or edX, which did not even exist a decade ago (Bonk, Lee, Reeves, & Reynolds, 2015; Haber, 2014).

Accepted: 23 April 2017
Correspondence: Richard E. Mayer, Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA 93106, USA. Email: mayer@psych.ucsb.edu

The case for using multimedia in eLearning

Multimedia instruction refers to presenting words and pictures that are intended to foster learning (Mayer, 2009, 2014a). The words can be in spoken form (such as narration) or in printed form (such as onscreen text). The pictures can be in static form (such as illustrations, diagrams, maps or photos) or dynamic form (such as animation or video). Some common uses of multimedia in e-learning include an animation, video or static graphics with accompanying narration; an animation, video or static graphic with accompanying onscreen text; or a computer-based interactive game, simulation or activity that includes spoken or printed text.

For hundreds of years, words have been the major format for instructional communications – progressing from spoken words (such as in lectures) to written words (such as on scrolls) to printed words (such as in books) to electronic words (such as on screens or via speakers). Verbal language is perhaps our species' single greatest educational technology, that is, our most useful tool for enhancing human learning and cognition. Its evolution into electronic form is a relatively new twist on using words for instruction, so the focus of this paper is on electronic instructional messages that involve both words and pictures.

The rationale for computer-based multimedia instruction is that people learn better from words and pictures than from words alone, which can be called the *multimedia principle* (Butcher, 2014; Mayer, 2009). In the case of e-learning, the multimedia principle posits that people learn better from computer-based words and pictures than from computer-based words alone. For example, suppose you listened to the following explanation of how a bicycle tire pump works three times: '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 through the hose.' When college students were tested on this lesson (narration only), they were able to generate about 3.2 correct answers on a problem-solving test (Mayer & Anderson, 1992).

However, consider what happens when we add a brief animation along with the verbal narration you are listening to, such as shown in Figure 1. When college students were tested on this lesson (narration–plus–animation), they were able to generate 7.7 correct answers on a problem-solving test – an improvement of more than 100% over the narration-only lesson. Interestingly, the groups did not differ significantly on

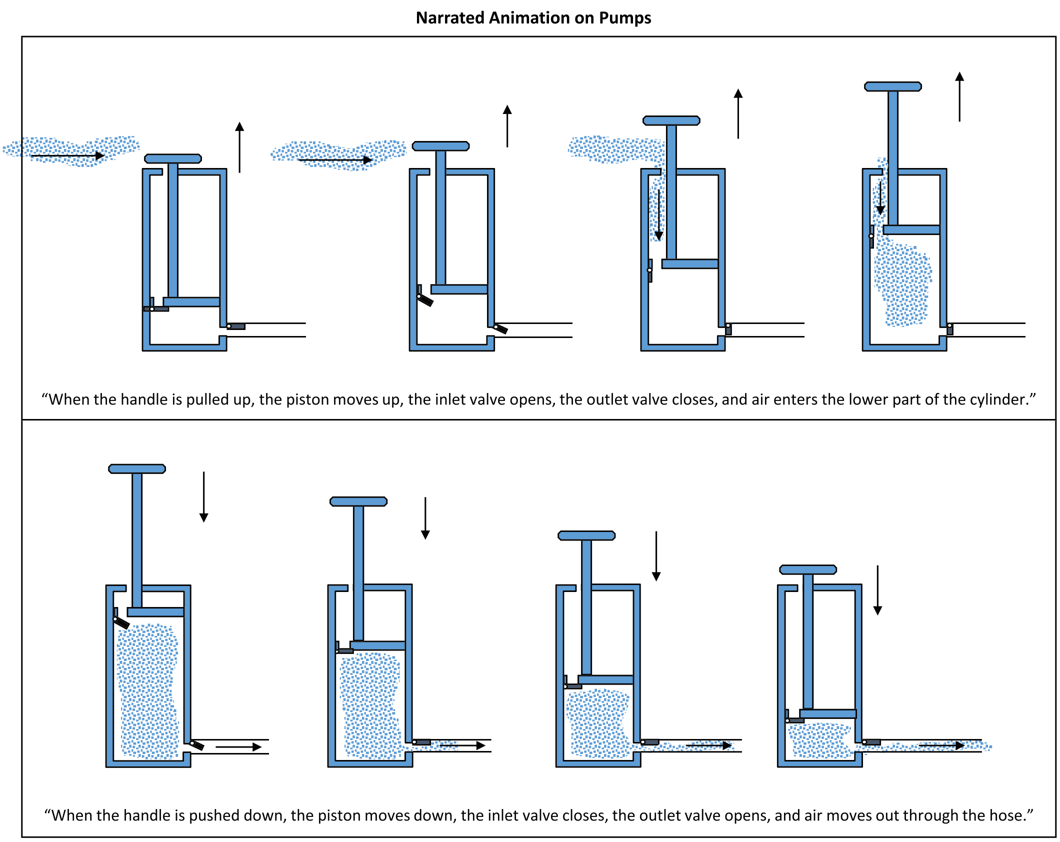

Figure 1 Narrated Animation on Pumps. [Colour figure can be viewed at wileyonlinelibrary.com]

recall of the narration – averaging between 70 and 80% on a recall test. In short, although adding pictures to words did not result in better remembering of the words, it did result in much better understanding as indicated by performance on a problem-solving transfer test. Overall, across five out of five experimental comparisons involving CAL about pumps, brakes, lightning and arithmetic (Mayer & Anderson, 1991, Expt. 2a; Mayer & Anderson, 1992, Expts. 1 & 2; Moreno & Mayer, 1999a, Expt. 1; Moreno & Mayer, 2002a, Expt. 1), students performed better when they learned with words and pictures than with words alone, yielding a median effect size of $d = 1.67$, which is a large effect. This kind of replicated support for the multimedia principle provides the rationale for using multimedia in e-learning. The multimedia principle is the founding design principle in this entire line of research. It provides a scientific basis for long-standing calls to incorporate graphics with instructional text dating back to the publication in 1658 of the first educational book to use illustrations and printed text, entitled *Orbis Pictus*, originally printed in 1658 (Comenius, 1887).

The case for using effective multimedia in e-learning

Although evidence for the multimedia principle highlights the potential value of adding graphics to text for improving student learning, not all uses of graphics are equally effective. In this section, I explore what research has to say about how to use words (such as spoken or printed text) and graphics (such as animation, video or static illustrations) to create effective multimedia instructional messages. I also explore the boundary conditions for each principle, including for whom, for which types of content and under what conditions each principle is most effective.

The theoretical basis for designing effective multimedia in e-learning is represented in the cognitive theory of multimedia learning, shown in Figure 2. Multimedia instruction enters the learner's information processing system through ears (leading to auditory sensory memory) and eyes (leading to visual sensory memory), where material is held for a very brief time. If the learner attends to some of the fleeting information in sensory memory (indicated by the *selecting* arrows), it can be transferred to working memory for further processing but must be actively processed within about 20 s to be kept in working memory. Next, the learner mentally organizes the incoming verbal information into a coherent verbal representation (or *verbal model*) and organizes the incoming visual information into a coherent pictorial representation (or *pictorial model*), as indicated by the *organizing* arrows. Finally, the learner activates relevant prior knowledge from long-term memory and brings it into working memory where it is integrated with the verbal and pictorial models, which are also integrated with each other (as indicated by the *integrating* arrow). The resulting learning outcome that is constructed in working memory can then be stored in long-term memory. This process can be iterative rather than strictly linear. Thus, meaningful learning depends on the learner engaging in active cognitive processing during learning including selecting, organizing and integrating.

As you can see, the model is based on three key principles from cognitive science: (a) *dual-channels principle* – people have separate information processing channels for visual and verbal information (Paivio, 1986); (b) *limited-capacity principle* – people can process only a few elements in each channel in working memory at any one time (Baddeley, 1999; Sweller, Ayres, & Kalyuga, 2011); and (c) *active processing principle* – meaningful learning occurs when people

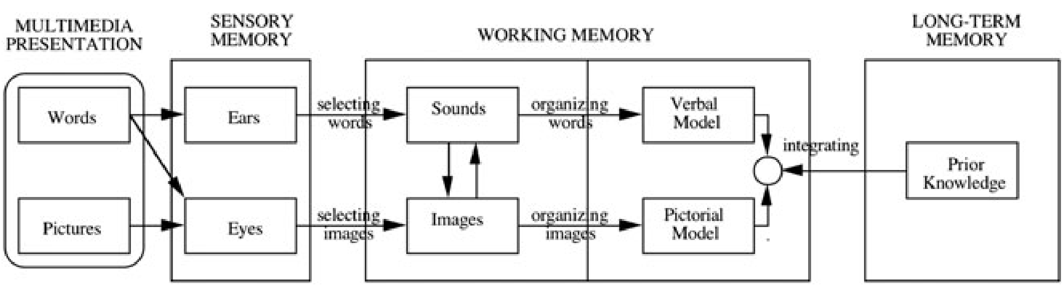

Figure 2 How Does Multimedia Learning Work? A Cognitive Theory of Multimedia Learning

engage in appropriate processing during learning, including selecting, organizing and integrating (Mayer, 2009; Wittrock, 1989).

For instructors, the major challenge is to create multimedia instructional messages that prime the three cognitive processes of selecting, organizing and integrating without overloading the visual and verbal channels in working memory. Given the limited capacity of working memory and the need to engage in active processing during learning, three major instructional goals are: (a) *reduce extraneous processing* – cognitive processing that does not support the instructional goal, which can be caused by poor instructional design; (b) *manage essential processing* – cognitive processing aimed at mentally representing the presented material, which can be caused by the complexity of the material for the learner; and (c) *foster generative processing* – cognitive processing aimed at making sense of the material, which can be caused the learner's motivation to exert effort to learn.

The following sections explore the research base for techniques that address each of these three goals, based on studies reviewed in *The Cambridge Handbook of Multimedia Learning* that deal only with computer-based lessons (Mayer, 2014b; Mayer & Fiorella, 2014; Mayer & Pilegard, 2014). The applicability of the findings to actual educational contexts is limited to the extent that many of the studies are short-term lab experiments with immediate tests.

Reducing extraneous processing: Coherence, signalling, redundancy, spatial contiguity and temporal contiguity principles

Table 1 lists five principles for reducing extraneous processing – coherence, signalling, redundancy, spatial contiguity and temporal contiguity principles. For each principle, the table also summarizes the number of experimental comparisons involving CAL that yielded positive results out of the total number of comparisons (under the *number* column) and the median effect size based on Cohen's *d* (under the *effect size* column). In each experimental comparison, we compare the performance on a test of learning outcome of a base group and a group that receives the same lesson with one instructional feature added. Each design principle in this section is intended to cause the learner to engage in less extraneous processing during learning, thereby freeing up cognitive capacity to engage in meaningful learning as represented by essential and generative processing.

First, consider the slides in Figure 3 from a computer-based slideshow lesson on how a virus causes a cold. The slide on the left contains a seductive detail – that is, interesting but irrelevant material – concerning research on the link between love-making and resistance to colds, whereas the slide on the right does not contain this extraneous material. You might expect that students would learn better from lessons that had a sprinkling of interesting (but irrelevant) facts about colds and virus on the grounds that students would be more interested. However, according to the cognitive theory of multimedia learning summarized in Figure 2, if people have to process extraneous material they will have less cognitive capacity available to processes the essential material in the lesson.

As predicted, the top row of Table 1 shows that in 12 out of 12 experimental comparisons involving computer-based lessons on lightning, brakes, ocean waves, ice age, cold viruses, biology and the history of distance learning (Mayer, DeLeeuw, & Ayres, 2007, Expts. 1 & 2; Mayer, Griffith, Jurkowitz, & Rothman, 2008, Expts. 1 & 2;

Table 1. Five research-based principles for reducing extraneous processing in e-learning

Principle	Description	Number	Effect size
Coherence	People learn better when extraneous material is excluded.	12 of 12	0.70
Signalling	People learn better when essential material is highlighted.	18 of 20	0.46
Redundancy	People learn better from graphics and narration that from graphics, narration, and on-screen text.	13 of 13	0.87
Spatial contiguity	People learn better when on-screen words are placed next to the corresponding part of the graphic.	8 of 8	0.79
Temporal contiguity	People learn better when corresponding narration and graphics are presented simultaneously.	8 of 8	1.30

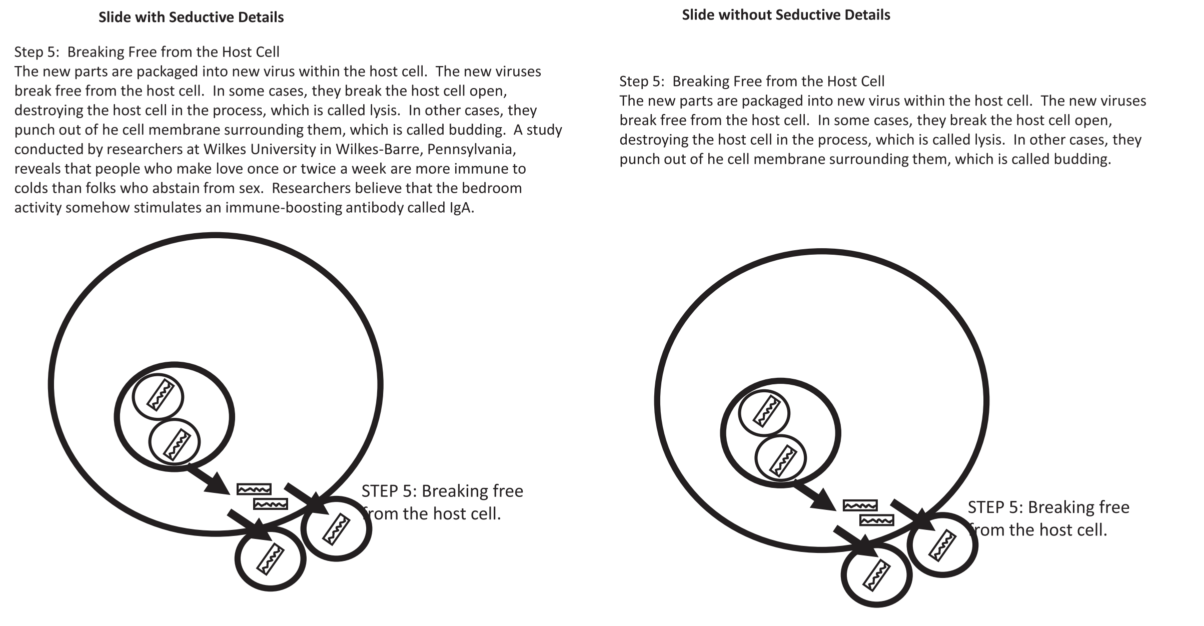

Figure 3 Does Adding Interesting Details Improve Learning?

Mayer, Heiser, & Lonn, 2001, Expt. 3; Mayer & Jackson, 2005, Expt. 2; Moreno & Mayer, 2000a, Expts. 1 & 2; Park, Moreno, Seufert, & Brunken, 2011; Sanchez & Wiley, 2006; Sung & Mayer, 2012), students scored better on transfer or comprehension tests when extraneous material was excluded rather than included, yielding a median effect size of d = 0.70, which is in the medium-to-large range. These findings support the *coherence principle*: People learn better from a computer-based multimedia lesson when extraneous material is excluded rather than included.

Concerning boundary conditions, adding emotionally interesting details to essential elements in a graphics on viruses (such as giving the virus and host cell human-like facial expressions) improved test performance (Mayer & Estrella, 2014; Plass, Heidig, Hayward, Homer, & Um, 2014; Um, Plass, Hayward, & Homer, 2012). Additional research is needed to determine when this kind of emotional design of graphics helps focus attention in relevant material and when it turns out to be distracting. In addition, there is emerging evidence that the coherence principle most strongly applies when the extraneous material is highly interesting (Mayer et al., 2008), when the learner has low working memory capacity (Rey, 2012; Sanchez & Wiley, 2006) and when the lesson is cognitively demanding (Park et al., 2011),

Second, consider the non-signaled script excerpt for a multimedia lesson on how airplanes achieve lift shown in the left side of Figure 4 and the signaled excerpt on the right (with bold words being stressed in speech). You might expect that students would learn better from the uncluttered version on the left which simply gives the explanation. However, if students need help in selecting the important information and organizing it, then using vocal emphasis to highlight the key terms and adding headings to show the organization may help guide the learner's cognitive processing during learning, consistent with the cognitive theory of multimedia learning. Signalling (also called visual cueing) can also involve highlighting relevant portions of the graphics using features such as arrows, colouring or spotlights.

As predicted, the second row of Table 1 shows that in 18 out of 20 experimental comparisons involving computer-based lessons on airplanes, visual perception, geography, mechanical systems, braking systems, cardiovascular system, teaching skills, piano mechanism, jet engine, neural networks, language production and fish locomotion (Amadieu, Marine, & Lemay, 2011; Boucheix, Lowe, Kemala-Putri, & Groff, 2013; Boucheix & Lowe, 2010, Expts. 1a & 1b; de Koning, Tabbers, Rikers, & Paas, 2007, 2010; Doolittle & Altstaedter, 2009; Jamet, Gavota, & Quaireau, 2008; Jarodzka, van Gog, Dorr, Scheiter, & Gerjets, 2013; Kriz & Hegarty, 2007; Mautone & Mayer, 2001, Expts. 3a & 3b; Mautone & Mayer, 2007; Moreno, 2007, Expts. 1 & 2; Naumann, Richter, Flender, Cristmann, & Groeben,

Excerpts from Non-signaled Script	Excerpts from Signaled Script
… surface on top of the wing is longer than on the bottom…	**Wing Shape: Curved Upper Surface is Longer** … surface on **top** of the wing is **longer** than on the **bottom**…
…air traveling over the curved top of the wing flows faster than air that flows under the bottom of the wing…	**Air Flow: Air Moves Faster Across Top of Wing** …air traveling over the curved **top** of the wing **flows faster** than air that flows under the **bottom** of the wing…
… the top surface of the wing now has less pressure exerted against it than the bottom surface of the wing…	**Air Pressure: Pressure on the Top is Less** … the **top** surface of the wing now has **less pressure** exerted against it than the **bottom** surface of the wing…

Figure 4 Does Adding Signalling Improve Learning?

2007, Expts. 1 & 2; Ozcelik, Arslan-Ari, & Cagiltay, 2010; Rey, 2010; Scheiter & Eitel, 2010), students scored better on transfer or comprehension tests when multimedia lessons contained signalling of the essential words or key features of the graphics, yielding a median effect size of d = 0.46, which is in the medium range. These findings support the *signalling principle*: People learn better from a computer-based multimedia lesson when essential parts of text or graphics are highlighted.

An important boundary condition is that the signalling principle was upheld for students with low prior knowledge but not for students with high prior knowledge (Naumann et al., 2007).

Third, as exemplified in Figure 5, should an online slideshow contain graphics and narration (as shown on the slide on the left) or graphics, narration and on-screen text that uses the words as the narration (as shown in the slide on the right)? You might suspect that having both spoken and printed text (as exemplified by the slide on the right) would be ideal so learners could choose the format that best suits their individual learning style. However, according to the cognitive theory of multimedia learning, presenting redundant spoken and printed text could cause an increase in extraneous processing, which interferes with learning. For example, learners may try to reconcile the stream of spoken words with stream of written words, rather than trying to understand what the words mean. Similarly, providing printed words as captions can split visual attention, so that learners cannot process the graphics if they are reading the caption. Thus, according to the cognitive theory of multimedia learning, presenting graphics and narration is most efficient because the graphics can be processed in the visual channel and the words can be processed in the verbal channel.

Consistent with these predictions, the third line of Table 1 shows that in 13 out of 13 experimental comparisons involving computer-based lessons on electrical engineering, lightning, botany and human memory (Austin, 2009, Expts. 1, 2, 3, & 4; Craig, Gholson, & Driscoll, 2002, Expt. 2; Jamet & Le Bohec, 2007; Kalyuga, Chandler, & Sweller, 1999, Expt. 1; Kalyuga, Chandler, & Sweller, 2000, Expt. 1; Mayer et al., 2001, Expts. 1 & 2; Moreno & Mayer, 2002a, Expts. 2a & 2b; Moreno & Mayer, 2002b, Expt. 2), students scored better on transfer or comprehension tests

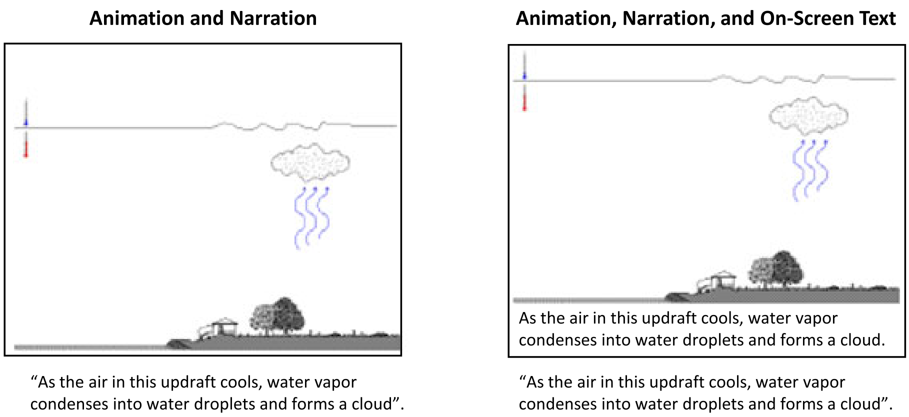

Figure 5 Which Instructional Method Leads to Better Learning from an Online Slideshow? [Colour figure can be viewed at wileyonlinelibrary.com]

when multimedia lessons contained graphics and narration rather than graphics, narration and on-screen text that was redundant with the narration, yielding a median effect size of $d = 0.87$, which is in the large range. These findings support the *redundancy principle*: People learn better from a computer-based multimedia lesson that contains graphics and narration rather than graphics, narration and on-screen text.

Concerning boundary conditions, Kalyuga et al. (1999, 2000) found a strong redundancy effect for students with low prior knowledge but not for students with high prior knowledge. Furthermore, adding spoken text to printed text has been shown to improve learning when there are no graphics (Adesope & Nesbit, 2012; Moreno & Mayer, 2002a), and adding redundant on-screen text to a narrated animation or narrated slideshow has been shown to improve learning when only one or two words are placed next to the part of graphic being discussed (Mayer & Johnson, 2008) or when the on-screen text is worded differently than the narration (Yue, Bjork, & Bjork, 2013).

Fourth, do you think people will learn better when printed words are presented as a caption below a graphic as shown in the slide on left side of Figure 6 or when printed words are placed near the corresponding part of the graphic as shown in the slide on the right side of Figure 6? You might think it would be a good idea to study the graphic first and then the words, or vice versa, in order to get a double exposure of the material, as with the separated presentation on the left slide. Also, you might favour the separated presentation because the learner has to engage in more cognitive activity to make connections between the words and graphics, creating what can be called a *desirable difficulty* (Bjork & Bjork, 2015). In contrast, the cognitive theory of multimedia learning predicts that separated presentation leads to more extraneous cognitive processing than does integrated presentation and therefore causes less cognitive capacity for cognitive processing needed for meaningful learning. When you have to scan back and forth between a caption and the relevant part of a graphic, this creates extraneous processing that wastes limited processing capacity. When the words are near the corresponding graphic, they signal where to look and thereby reduce unproductive scanning.

Consistent with this prediction, the fourth line of Table 1 shows that in eight out of eight experimental comparisons involving computer-based multimedia lessons on lightning, pumps, brakes, electrical circuits, statistics, medical procedures and how kidneys work (Austin, 2009; Bodemer, Ploetzner, Feuerlein, & Spada, 2004, Expt. 2; Cierniak, Scheiter, & Gerjets, 2009; Johnson & Mayer, 2012, Expts. 1, 2, & 3; Kester, Kirschner, & van Merrienboer, 2005; Moreno & Mayer, 1999b, Expt. 1; Pociask & Morrison, 2008), students learned better with integrated multimedia lessons than separated multimedia lessons, yielding a median effect size of d = 0.79, which is in the large range. In these studies, the presented graphics and on-screen text were identical in the integrated and separated conditions, with the only difference being whether the words were placed near or far from corresponding parts of the graphic. These findings, also consistent with a review by Ginns (2006), support the *spatial contiguity principle*: People

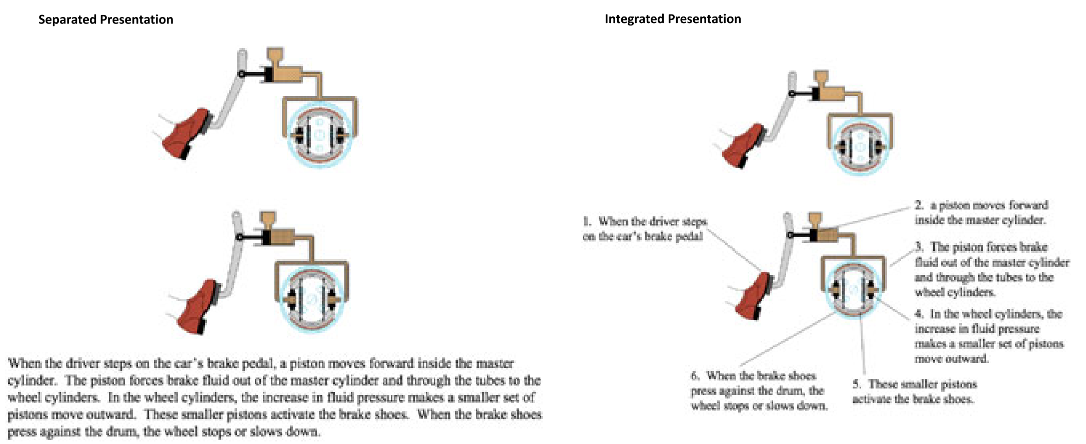

Figure 6 Which Instructional Method Leads to Better Learning? [Colour figure can be viewed at wileyonlinelibrary.com]

learn better from multimedia lessons when printed words are placed near rather than far from corresponding parts of the graphic in the screen.

The fifth way to reduce extraneous processing involves the temporal ordering of narration and graphics. Consider two ways to present corresponding narration and graphics, simultaneously so that spoken words correspond to what is being shown in the screen (in the right side of Figure 7) or successively so that the narration goes before or after the slideshow or animation (in the left side of Figure 7). You might suppose that successive presentation of words then graphics (or vice versa) would be best because it involves two separate exposures to the material and takes twice as much time as simultaneous presentation. In contrast, however, the cognitive theory of multimedia learning predicts that successive presentation will create massive extraneous processing because you have to keep the entire narration in working memory until the graphics are presented, or vice versa, leaving less cognitive capacity for making sense of the incoming material. With simultaneous presentation learners can have corresponding words and pictures in their working memory at the same time, thereby facilitating the cognitive process of integrating, which is instrumental for meaningful learning.

This interpretation is upheld by the results summarized in the fifth line of Table 1, showing that in eight of eight experimental comparisons, involving computer-based multimedia lessons on pumps, brakes, lightning and how the lungs work (Mayer & Anderson, 1991, Expts. 1 & 2; Mayer & Anderson, 1992, Expts. 1 & 2; Mayer, Moreno, Boire, & Vagge, 1999, Expts. 1 & 2; Mayer & Sims, 1994), students performed better on transfer tests when narration and graphics were presented simultaneously rather than successively, yielding a median effect size of $d = 1.30$, which is a large effect. This pattern of results, also consistent with a meta-analysis by Ginns (2006), supports the temporal continuity principle: People learn better from multimedia lessons when narration and graphics are presented simultaneously rather than successively.

Successive
First listen to narration
Then view slides

OR

First view slides
Then listen to narration

Simultaneous
Receive narrated slides

Figure 7 Which Instructional Method Leads to Better Learning?

Concerning boundary conditions, the temporal contiguity principle does not apply when the material is presented in small chunks (Mayer et al., 1999; Moreno & Mayer, 1999b) or when the material is simple for the learner (Ginns, 2006).

Overall, Table 1 lists five techniques for reducing extraneous processing (and thereby facilitate learning) that have substantial empirical support (Mayer & Fiorella, 2014).

Managing essential processing: Segmenting, pre-training and modality principles

The goal of the design principles in the previous section was to reduce extraneous processing, but even if we greatly reduce extraneous processing, sometimes a computer-based multimedia lesson may be so complex for a learner that the required essential processing exceeds the learner's cognitive capacity. In this situation, we need design principles for managing essential processing, which I address in this section. As summarized in Table 2, three ways to help learners manage how they mentally represent complex material are by breaking a computer-based multimedia lesson into self-paced segments (segmenting principle), familiarizing learners with the key terms before the lesson (pre-training principle) and presenting words in spoken form (modality principle).

As an example of the segmenting principle, consider two versions on a multimedia lesson shown in Figure 8. On the left is a frame from a continuous multimedia lesson consisting of animation with concurrent narration, and on the right is a frame from the end of a short segment – approximately 10 s – which has a CONTINUE button that the learner can click in order to go on to the next short segment. Which version will be more effective – a continuous presentation (in this case lasting about 2.5 min) or a segmented presentation (in this case with each segment conveying one major event and lasting about 10 s). You might think the segmented presentation would be a poor choice because having to keep pressing the CONTINUE button would be disruptive. However, in contrast, the cognitive theory of multimedia learning favours the segmented presentation on the grounds that it allows the learner to fully digest a chunk of information – including

Table 2. Three research-based Principles for managing essential processing in e-learning

Principle	Description	Number	Effect size
Segmenting	People learn better when a multimedia lesson is presented in small user-paced segments.	12 of 12	0.70
Pre-training	People learn when they learn the key terms prior to receiving a multimedia lesson.	18 of 20	0.46
Modality	People learn better from a multimedia lesson when words are presented in spoken form.	42 of 51	0.72

integrating the corresponding words and pictures – before moving on to the next one.

This prediction is supported in the first line of Table 2, showing that in 10 of 10 experimental comparisons, involving computer-based multimedia lessons on lightning, electric motors, pulley systems, how the human eye works, astronomy, history and teaching skills, students who learned with segmented lessons (Boucheix & Schneider, 2009; Hasler, Kersten, & Sweller, 2007; Hassanabadi, Robatjazi, & Savoji, 2011; Lusk et al., 2009; Mayer & Chandler, 2001, Expt. 2; Mayer, Dow, & Mayer, 2003a, Expts. 2a & 2b; Moreno, 2007, Expts. 1 & 2; Stiller, Freitag, Zinnbauer, & Freitag, 2009) performed better on comprehension or transfer tests than students who learned with continuous lessons covering the identical material, with a median effect size of $d = 0.77$, which is in the large range. This pattern of results supports the *segmenting principle*: People learn better from multimedia lessons that are broken into self-paced segments.

A potential boundary condition is that the segmenting principle may apply for learners with low working memory capacity but not high working memory capacity (Lusk et al., 2009).

Next, suppose you were about to receive a computer-based narrated animation explaining how a car's braking system works that lasts about 1 min. The, narrated animation tells you about how a change in one part of the system causes a change in another part (e.g., a piston moves forward in the master cylinder causes an increase in fluid pressure) and so on. Do you think it would help you to receive a brief pre-training in the names and locations of the key components in the braking system (such as 'piston in master cylinder')? For example, you could click on any part of the drawing system, and the computer will tell you the name and characteristics of the part, as exemplified in Figure 9. According to the cognitive theory of multimedia learning, being familiar with the terms used in the narrated animation subsequently can help you focus on learning the cause-

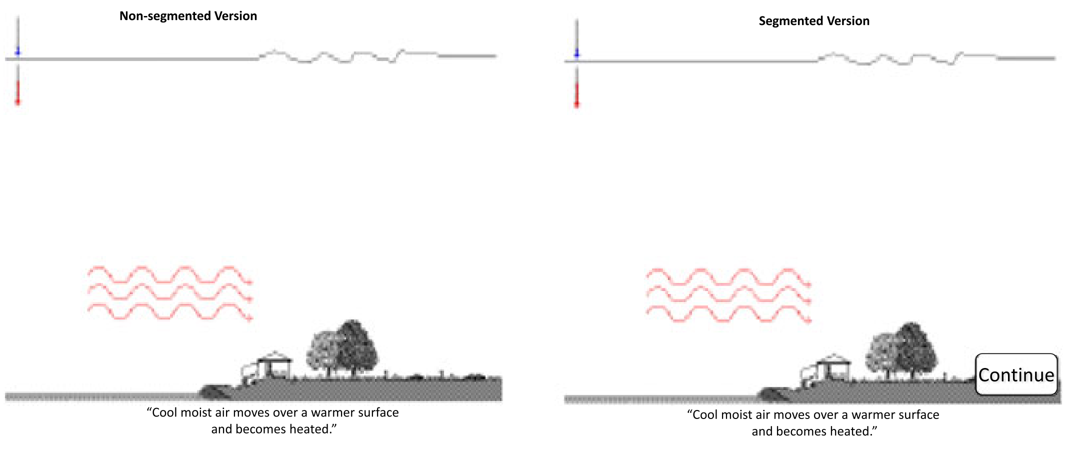

Figure 8 Do People Learn Better When a Continue Button Is Added After Each Segment? [Colour figure can be viewed at wileyonlinelibrary.com]

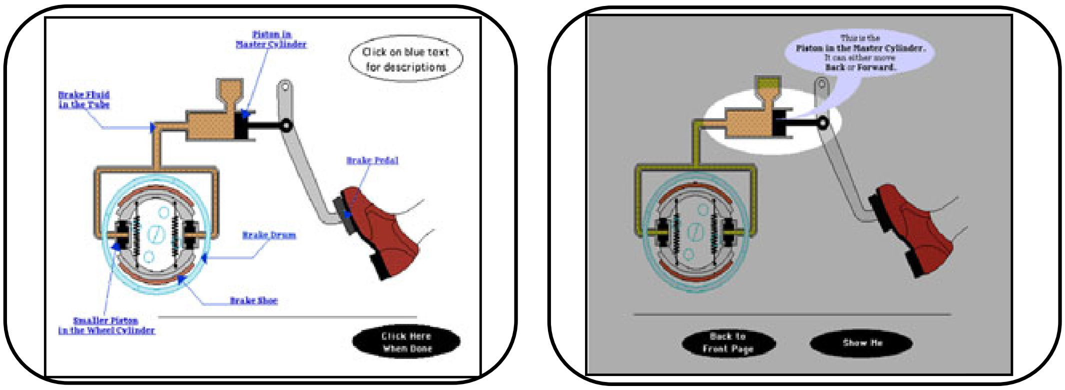

Figure 9 Do People Learn When They Receive Pretraining in the Names and Characteristics of Key Parts? [Colour figure can be viewed at wileyonlinelibrary.com]

and-effect relations in the system, which should result in better learning.

As expected, the second line of Table 2 shows that in nine out of ten experimental comparisons, involving computer-based multimedia lessons on brakes, pumps, geology, statistics, electricity, neural networks and pulley systems (Eitel, Scheiter, & Schüler, 2013; Kester, Kirschner, & van Merrienboer, 2004a, 2004b, 2006a; Kester, Lehnen, van Gerven, & Kirschner, 2006b; Mayer, Mathias, & Wetzell, 2002a, Expts. 1, 2, & 3; Mayer, Mautone, & Prothero, 2002b, Expt. 3), students scored higher on posttests when they received pre-training, yielding a median effect size of $d = 0.75$, which is in the high range. This pattern of results supports the *pre-training principle*: People learn better from computer-based multimedia lessons when they receive pre-training in the key elements.

Concerning boundary conditions, more work is needed to determine how to design effective pre-training for lessons that teach how to solve problems via interactive simulations (Kester, Kirschner, & van Merrienboer, 2004a, 2004b, 2006a; Kester, Lehnen, van Gerven, & Kirschner, 2006b).

The third line in Table 2 examines the issue of whether words in a multimedia lesson should be spoken or printed, as exemplified in the left and right panels of Figure 10, respectively. You might think that would not matter whether words are spoken or printed because they convey the same information. However, according to the cognitive theory of multimedia learning, the visual

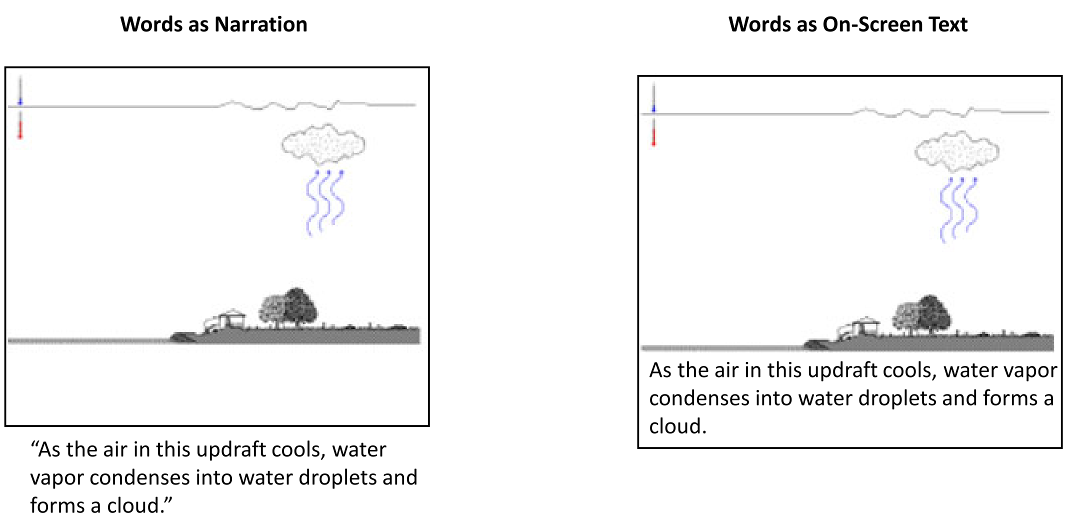

Figure 10 Which Instructional Method Leads to Better Learning? [Colour figure can be viewed at wileyonlinelibrary.com]

channel could become overloaded when words are presented as on-screen text in a caption. Receiving graphics and printed words on the screen can create a split attention situation for learners, in which they cannot read the words if they are looking at the graphic and they cannot look at the graphic if they are reading the words. This potential overload situation can be mitigated by offloading the words into the verbal channel so the graphics enter your cognitive system through your eyes and are processed in your visual system whereas the words enter through your ears and are processed in your verbal system. In this way, some of the demands on the visual channel are offloaded onto the verbal channel, thereby making the best use of the available processing capacity in both channels.

Consistent with this theoretical interpretation, students tend to learn better from computer-based multimedia lessons when words are spoken rather printed on the screen, yielding a median effect size of $d = 0.72$, which is in the medium-to-large range. Overall, in 42 out of 51 experimental comparisons, spoken words resulted in better learning outcome test performance than having the same words printed on screen, including computer-based multimedia lessons on topics such as lightning, brakes, botany, math problems, electrical engineering, electric motors, biology, instructional design, probability, computer networking, human speech and tornados (Atkinson, 2002, Expts. 1a, 1b, & 2; Cheon, Crooks, & Chung, 2014; Craig et al., 2002, Expt. 2; Crooks, Cheon, Inan, Ari, & Flores, 2012; Harskamp, Mayer, & Suhre, 2007, Expts. 1 & 2a; Jeung, Chandler, & Sweller, 1997, Expts. 1, 2, & 3; Kalyuga et al., 2000, Expt. 1; Kühl, Scheiter, Gerjets, & Edelmann, 2011; Leahy & Sweller, 2011; Lindow et al., 2011, Expt. 2; Mayer, Dow, & Mayer, 2003a, Expt. 1; Mayer & Moreno, 1998, Expts, 1 & 2; Mayrath, Nihalani, & Robinson, 2011, Expts. 1 & 2; Moreno, Mayer, Spires, & Lester, 2001, Expts. 1, 4a, 4b, 5a, & 5b; Moreno & Mayer, 2002b, Expts. 1a, 1b, 1c, 2a, & 2b; O'Neil et al., 2000, Expt. 1; Schuler, Scheiter, Rummer, & Gerjets, 2012; Schuler, Scheiter, & Gerjects, 2013, Expts. 1 & 2; Schmidt-Weigand, Kohnert, & Glowalla, 2010a, Expts. 1a, 1b, & 1c; Schmidt-Weigand, Kohnert, & Glowalla, 2010b, Expts. 1a & 1b; Tabbers, Martens, & van Merrienboer, 2004; Witteman & Segers, 2010; Wong, Leahy, Marcus, & Sweller, 2012; Wouters, Paas, & Van Merriënboer, 2009). This pattern allows me to propose the *modality principle*: People learn better from computer-based multimedia lessons when words are spoken rather than printed. Similarly, in a meta-analysis of research on the modality principle, Ginns (2005) reported a weighted mean effect size of $d = 0.43$ favouring spoken words over printed words in multimedia lessons.

The modality principle is the most studied of all the multimedia design principles, which allows us to better pinpoint boundary conditions. In a recent review, Reinwein (2012) found evidence that the modality principle may be strongest for system-paced rather than learner-paced lessons, dynamic rather than static graphics and transfer rather than retention tests. This pattern is consistent with the cognitive theory of multimedia to the extent that off-loading demands on the visual channel was most effective for situations that involved visually demanding lessons – that is, fast-paced lessons and with lots of motion. Similarly, the modality principle tends to apply more strongly when the learners have a low rather than high level of prior knowledge, when verbal segments are short rather than long and when words are familiar rather than unfamiliar (Mayer & Pilegard, 2014). These patterns are also consistent with the cognitive theory of multimedia learning. Thus, on-screen text is likely to be more effective when the text is so long that the reader may not be able to hold it all in working memory, when the text contains words that are unfamiliar to the learner, when the learner is not a native speaker of the language used in the lesson and when the reader has problems with hearing. In short, research is helping to pinpoint the conditions under which the modality principle does and does not apply.

Overall, there is substantial and growing research support that segmenting, pre-training and modality principles can be effective in helping learners manage essential processing – that is, helping them handle the task of representing material that is complicated for them (Mayer & Pilegard, 2014).

Fostering generative processing: Multimedia principle, personalization, voice and embodiment principles

Even if we design lessons that help reduce extraneous processing and manage essential processing, as explored in the previous two sections, learners might not be motivated to use their available cognitive capacity to make sense of the presented material. In this section, I explore multimedia design principles aimed at fostering

Table 3. Three research-based principles for fostering generative processing in e-learning

Principle	Description	Number	Effect size
Personalization	People learn better when the words in a multimedia lesson are presented in conversational style rather than formal style.	14 of 17	0.79
Voice	People learn better from a human voice than a machine-like voice.	5 of 6	0.74
Embodiment	People learn better when an onscreen agent uses human-like gestures and movement.	13 of 14	0.36

generative processing – that is, cognitive processing aimed at understanding the material more deeply. In short, we want to prime learners to exert effort to make sense of the material. Table 3 lists three multimedia design principles intended to foster generative processing in learners: personalization principle, voice principle and embodiment principle. For each, I also summarize the research evidence including the median effect size.

The main theme in this section is that social cues in the lesson can prime a social stance in the learner, in which the learner sees the instructor as a social partner. When the learner feels social partnership with the instructor, the learner will exert more effort to understand what the instructor is saying, which results in better learning outcomes. This is the causal chain proposed in social agency theory (Mayer, 2014b), and the basis for considering three social cues for on-screen agents intended to prime a social response in the learner – using conversational language, using human voice and using human-like gestures.

First, Figure 11 shows portions of a script from a narrated animation on how the human respiratory system works, with the script on the left using formal style (i.e., non-personalized) and the script on the right using conversational style by changing ‘the’ to ‘your’ throughout the lesson (i.e., personalized). You might think both lessons would be equally effective because they present the same information about the human respiratory system. However, social agency theory and the cognitive theory of multimedia learning predict that learners will be more motivated to understand the personalized version and therefore perform better on transfer and comprehension tests.

The first line in Table 3 provides evidence for this prediction, by showing that in 14 of 17 experimental comparisons, students learned better from computer-based multimedia lessons that contained conversational language rather than formal language, yielding a median effect size of $d = 0.79$, which is a large effect. The studies involved lessons on the human respiratory system, lightning, botany, astronomy, chemistry and engineering, and some studies compared polite versus direct wording (Kartal, 2010; Mayer, Fennell, Farmer, & Campbell, 2004, Expts. 1, 2, & 3; McLaren et al., 2007; McLaren, DeLeeuw, & Mayer, 2011a, 2011b; Moreno & Mayer, 2000b, Expts., 1, 2, 3, 4, & 5; Moreno & Mayer, 2004, Expts. 1a & 1b; Wang et al., 2008). This pattern of results provides support for the *personalization principle*: People learn better from computer-based multimedia lessons when the words are presented in conversational style rather than formal style. In general, the personalization principle can be implemented by using first and second person constructions (e.g., using ‘I’ and ‘we’ and ‘you’) either with voice or on-screen text.

An important boundary condition in studies comparing polite versus direct wording is that the personalization principle is found for inexperienced or low-prior knowledge learners but not for experienced or high-prior knowledge learners. In another meta-analysis on the personalization principle, Ginns, Martin, and Marsh (2013) reported an effect size of $d = 0.54$ and noted that the personalization works best for short lessons (i.e., less 35 min) rather than long lessons. These boundary conditions suggest that personalization is most important in the early stages of building rapport with a computer-based instructor.

Non-Personalized
“During inhaling, the diaphragm moves down creating more space for the lungs, air enters through the nose or mouth, moves down through the throat and bronchial tubes to tiny air sacs in the lungs…”

Personalized
“During inhaling, your diaphragm moves down creating more space for your lungs, air enters through your nose or mouth, moves down through your throat and bronchial tubes to tiny air sacs in your lungs…”

Figure 11 Which Wording Style Leads to Better Learning?

Second, consider a computer-based lesson in which an on-screen character explains how solar cells work while standing next to a series of slides. Would it matter whether the charter's voice was a recording of a human voice or a machine synthesized voice that can easily be discerned? You might expect both versions would yield similar results because both present exactly the same information. However, according to social agency theory and the cognitive theory of multimedia learning, the human voice is more likely to prime a social connection whereas the machine-like voice is likely to break the illusion of a social partnership. Thus, people should learn more deeply when the lesson includes a human voice rather than a machine voice.

The second line of Table 3 is consistent with these predictions, by showing that in five of six experimental comparisons, people learn better from a computer-based narrated lesson when the voice is human rather than machine-like, yielding a median effect size of $d = 0.74$, which is close to the large range. This includes lessons on solar cells, lightning and math problems (Atkinson, Mayer, & Merrill, 2005, Expts. 1 & 2; Mayer & DaPra, 2012, Expts. 2a & 2b; Mayer, Sobko, & Mautone, 2003b, Expts. 1 & 2). Overall, this pattern of results yields the *voice principle*: People learn better from computer-based multimedia lessons when words are spoken in a human voice rather than a machine-like voice.

An important boundary condition is that the voice principle does not work well when there are other cues in the lesson that disrupt the illusion of social partnership, such as when an onscreen agent does not use human-like gesture and facial expression.

Finally, suppose you are watching a video of a lecture on the Doppler effect in which the instructor explains while standing next to a white board with corresponding drawings on it (as shown in the left side of Figure 12). Contrast this with watching a video of the same lecture with the same script and the same drawings, except the instructor draws the illustrations on the whiteboard as she explains (as shown in the right side of Figure 12). You might predict that you would learn equally well from the two versions because both present the same information. However, according to social agency theory and the cognitive theory of multimedia learning, seeing someone draw as she explains could prime more of a social connection in the learner in which the learner mentally participates in the act of drawing along with the instructor, leading to deeper learning. In short, people might try harder when they see a high-embodied on-screen agent – one that draws while talking or employs gesture, facial expression and eye gaze – rather than a low-embodied on-screen agent – one who stands motionless, in contrast to how humans usually behave.

Consistent with social agency theory, the third line of Table 3 shows that in 13 of 14 experimental comparisons, people learned better from a computer-based multimedia lesson – including animations and videos – when the onscreen agent used human-like gestures and movements while talking, yielding a median effect size of $d = 0.36$, which is in the small-to-medium range. This includes lessons on the Doppler effect, solar cells, astronomy, electricity, math problems, word-processing software and the human heart (Baylor & Kim, 2009; Dunsworth & Atkinson, 2007; Fiorella & Mayer, 2016, Expts, 1 & 2; Frechette & Moreno, 2010; Lusk & Atkinson, 2007; Mayer & DaPra, 2012, Expts, 1, 2a, 2b, 3a, & 3b; Moreno, Reislein, & Ozogul, 2010). The results support the *embodiment principle*: People learn better from

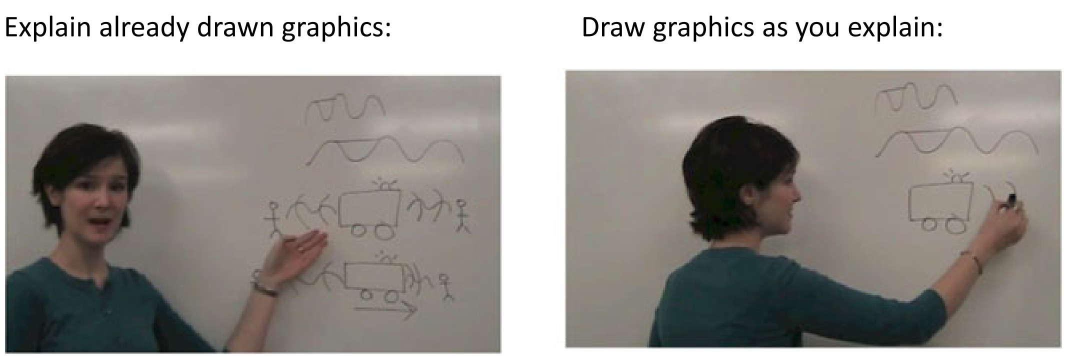

Figure 12 Which Method Leads to Better Learning from a Video Lecture? [Colour figure can be viewed at wileyonlinelibrary.com]

computer-based multimedia lessons that contain high-embodied on-screen agents rather than low-embodied on-screen agents.

A possible boundary condition is that the embodiment effect may not apply when the learners have high prior knowledge or when the lesson has negative social cues such as machine-like voice (Mayer, 2014b).

The future of using multimedia in eLearning

This review shows that over the past three decades evidence has been accumulating concerning how to design computer-based learning experiences (or e-learning). This paper presents a continuation of my efforts to systematize what we know about how to design effective e-learning into a set of research-based principles (Mayer, 2009, 2014a). Consistent with Shavelson and Towne's (2002) advice that replication is a fundamental requirement for scientific research in education, I focus on principles that have been tested across numerous situations. In the future, replication studies will be needed to determine the robustness of the design principles proposed in this paper. In particular, as explained in the research agenda in the following section, it would be useful to examine e-learning in more realistic learning environments and on delayed tests.

Research on instructional design principles for e-learning satisfies both practical and theoretical goals. On the practical side, the development of evidence-based principles gives us something useful to share with instructional designers, although with caution and humility given the need for more studies on long-term effects and in realistic learning environments. On the theoretical side, research on e-learning helps strengthen cognitive theories of learning such as the cognitive theory of multimedia learning (Mayer, 2009) and cognitive load theory (Sweller et al., 2011). In the future, more fine-grained research is needed to help provide more detail about the cognitive mechanisms involved in e-learning, as elaborated in the following section in a research agenda.

A sign of the growing maturity of the field of instructional design for e-learning is the discovery of boundary conditions for when principles do and do not apply, and the ability to explain these boundary conditions in terms of cognitive theories of academic learning. Future research is needed to pinpoint and replicate when the principles apply including for which kinds of learners, which kinds of learning objectives and which kinds of content areas.

Although most of the experiments reviewed for this paper involved instructional lessons such as narrated animations or slideshows, research in the future is likely to explore a wider set of venues including instructional video, educational games, virtual reality and mobile devices. It is important to determine whether the principles apply when instruction is delivered via new media. There is no reason to suspect that people's cognitive systems change depending on the instructional medium they are using, so I suspect that the same design principles will apply with new instructional technologies. However, new technologies will afford new instructional techniques that can help add new principles.

A research agenda

Based on this brief overview of research on design principles for using multimedia in e-learning, I offer a research agenda for future work.

Replicate with new materials and learners

Although some journal editors and reviewers have downplayed the value of replication studies as not making a significant contribution, there is growing consensus that replication is fundamental to progress in our field. In their classic National Research Council report, *Scientific Research in Education*, Shavelson and Towne (2002, p. 4) list 'Replicate and generalize across studies' as one of the six basic principles of educational research. Replication studies are particularly needed in light of the growing importance of meta-analyses, which combine studies that all examine the same basic effect (Hattie, 2009), as well as for contributing to the ongoing conversation about the so-called 'crisis of replication' in which some oft-cited effects in psychology do not appear to replicate (Pashler & Wagenmakers, 2012, p. 528; Stroebe & Strack, 2014, p. 59). Thus, this research agenda is consistent with the widening recognition in our field that original research journals should publish methodologically-sound replication studies – perhaps as brief reports – including both those that confirm and disconfirm earlier findings. In short, replication studies deserve a valued place in our field as they contribute to the scientific research base for building theory and making practical recommendations.

Replicate with new media
Early work on multimedia learning tended to involve paper-based materials and shifted somewhat to computer-based venues such as narrated animation on desktop computers. Continued work is needed that explores the efficacy of design principles with other media and platforms such as video, serious games, virtual reality, tablets and smartphones.

Replicate with longer lessons
Much of the published work involves short-term studies using single lessons that last less than an hour, so it is worthwhile to examine to the robustness of design principles across multiple lessons totaling 5 to 10 h or more. Such work is needed to overcome the concern that principles based on short-term studies may not apply to long-term learning.

Replicate in authentic settings
Much of the published work involves laboratory studies, so it is worthwhile to examine the robustness of design principles for lessons for students in real classes. Such work is needed to overcome the concern that principles based on lab studies may not apply to actual educational venues.

Replicate with delayed tests
Much of the published work involves immediate tests, but some researchers have shown that delayed tests can be a better measure of learning (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). Thus, it is worthwhile to examine the robustness of design principles when delayed tests are administered.

Replicate with multi-leveled tests
Some of the published work involves only retention tests, which measure the amount of information the learner can remember, but deep learning is best measured using transfer tests, which require the learner to use what is learned to answer new questions (Anderson et al., 2001). Specifically, I recommend a battery of posttests that includes both retention and transfer tests.

Examine boundary conditions by varying type of material, type of learner, type of media, length of lesson, timing of test, type of test item and other potentially relevant factors
Although many of the founding studies in multimedia learning simply compared treatment and control groups, there is a need to determine the strength of treatment effects across potentially moderating factors. This requires factorial designs in which one factor is the presence or absence of the design feature and the other factor is potentially relevant moderating factor such as low versus high prior knowledge learners or desktop versus immersive virtual reality venues. The goal is to determine boundary conditions for each principle, that is, the conditions under which it produces strong effects or not.

Examine new techniques for reducing extraneous processing.
Although this review shows that the literature is ripe with several well-established principles for how to reduce extraneous processing, there is a need to search for new techniques and extensions, such as determining the minimal amount of verbal and visual information that is needed.

Examine new techniques for managing essential processing
Although this review shows that the literature is ripe with several well-established principles for how to manage essential processing, there is a need to search for new techniques and extensions, such as how best to segment the material on a PowerPoint slide.

Examine new techniques for fostering generative processing
Although this review shows that the literature yields several well-established principles for how to foster generative processing, there is a need to search for new techniques and extensions, including a better understanding of role of social cues such as watching videos of instructors draw as they explain.

Examine the benefits of combining design principles
Much of the previous work uses a value added approach that examines the effectiveness of adding one feature to a base lesson. A useful direction for future work is to examine the added value of adding several features. In particular, it would be useful to know whether the effects of design principles can be additive, which would allow for creating maximally effective lessons.

Examine the role of motivation in multimedia learning
Although the crucial role of motivation in academic learning is widely recognized (Wentzel & Miele, 2016),

research is needed that incorporates the role of motivation in theories concerning the design of multimedia instruction, such as begun with Moreno's cognitive affective theory of learning with media (Moreno & Mayer, 2007). Motivation is an understudied aspect of multimedia learning that needs to better addressed in future research.

Examine the role of metacognition in multimedia learning
Another understudied aspect of multimedia learning concerns the role of metacognition, including the examining the role of learner's executive function skills such as shifting, updating and inhibiting (Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000).

Include direct measures of cognitive processing during learning
Although a fundamental component of multimedia research involves using posttests given after learning to assess learning outcomes, an understudied aspect of multimedia learning concerns measuring cognitive processing. Although self-report surveys of mental effort and perceived difficulty are commonly used in today's research, they are subject to the perils inherent in self-reports and need to be supplemented or replaced with direct measures of cognitive processing during learning (i.e., behavioural or physiological measures). Being able to operationalize *cognitive load* using direct measures is a major task on the research agenda (Brunken, Seufert, & Paas, 2010). Promising techniques for measuring processing during learn include using dual-task methodologies, eye-tracking, brain measures such as EEG and fMRI, and other physiological measures such as heartrate, blood flow and galvanic skin response.

Develop instruction in strategies for effective learning with multimedia materials and for creating effective multimedia materials
The increase in multimedia communications requires a reconceptualization of literacy to include *multimedia literacy*, which includes the ability to learn from presentations involving both words and graphics as well as the ability to produce messages involving both words and graphics that effectively communicate with others (Mayer, 2008). Students currently receive extensive training in reading comprehension strategies for reading text and in writing strategies for producing written essays, but an important next step on the research agenda is to study how best to teach students how comprehend and create messages involving graphics and words.

Expand and explicate theoretical accounts of multimedia learning
The existing theories of multimedia learning, such as the cognitive theory of multimedia learning (Mayer, 2009), cognitive load theory (Sweller et al., 2011) and the integrated model of text and picture comprehension (Schnotz, 2014), represent a start based on the available research evidence, but more work is needed in specifying the mechanisms of learning and in incorporating the role of factors such as motivation and metacognition.

Improve the methodological rigor of research on multimedia learning
The larger field of educational research is beset by arguments about which 'ism' should dominate (e.g., behaviourism, cognitivism, constructivism, social constructivist, situated constructivism, etc.), whether researchers should use qualitative or quantitative methodologies, and whether we should ease the scientific requirements for experimental research in schools. This research agenda calls for an end to the fruitless search for the best 'ism,' which has shown itself to be a fruitless distraction for our field. It also calls for an end to the *paradigm wars* that pit qualitative and quantitative methods or experimental versus observational methods against one another, and instead calls for choosing methods (often in combination) that best suit the research question under investigation (Shavelson & Towne, 2002). Finally, when experimental methods are used (such as to test the effectiveness of instructional interventions), in order to be able to interpret the results, researchers should adhere to the fundamental standards of experimental control (e.g., including a control group rather than simply testing for pretest-to-posttest gains), random assignment (e.g., randomly assigning students to treatments) and appropriate measures of learning (e.g., including measures of learning outcome rather than solely focusing on self-reports). A prerequisite for progress in the research in multimedia learning is that the research base contains findings gleaned from methodologically sound studies.

Understand the context of implementation
Finally, research is needed to examine the understudied issue of how instructional interventions – such as the multimedia principles – are affected by the context in which they are implemented including the social, organizational, collaborative and cultural context.

The take-home message from this overview of research on multimedia learning is that we know a lot about how to design multimedia e-learning, based on a growing research base and cognitive theories of learning, but continued research is needed to address the 18 points in the research agenda. Overall, from my vantage point in educational psychology, the future of e-learning looks bright to the extent that we continue to take a scientific approach, in which educational decisions are informed by research evidence. In the words of E. L. Thorndike (1906, p. 206), written more than a century ago, we still seek a future in which 'leaders in education direct their choices of methods by the results of scientific investigation rather than general opinion.'

Acknowledgements

Preparation of this paper was supported by grant N00014162046 from the Office of Naval Research.

Conflict of interest

No conflicts of interest have been declared.

References

Adesope, O. O., & Nesbit, J. C. (2012). Verbal redundancy in multimedia learning environments: A meta-analysis. *Journal of Educational Psychology*, *104*, 250–263.

Amadieu, F., Marine, C., & Lemay, C. (2011). The attention guiding effect and cognitive load in the comprehension of animations. *Computers in Human Behavior*, *27*, 36–40.

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

ATD (2014). *State of the industry report 2014*. Alexandria, VA: Association for Talent Development.

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology*, *94*, 416–427.

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.

Austin, K. A. (2009). Multimedia learning: Cognitive individual differences and display design techniques predict transfer learning with multimedia learning modules. *Computers and Education*, *53*, 1339–1354.

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

Baylor, A. L., & Kim, S. (2009). Designing nonverbal communication for pedagogical agents: When less is more. *Computers in Human Behavior*, *25*, 450–457.

Bjork, E. L., & Bjork, R. A. (2015). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York: Worth.

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and Instruction*, *14*, 325–341.

Bonk, C. J., Lee, M. M., Reeves, T. C., & Reynolds, T. H. (2015). *MOOCs and open education around the world*. New York: Routledge.

Boucheix, J. M., & Lowe, R. K. (2010). An eye-tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction*, *20*, 123–135.

Boucheix, J. M., Lowe, R. K., Kemala-Putri, D., & Groff, J. (2013). Cueing animations: Dynamic signaling aids information extraction and comprehension. *Learning and Instruction*, *25*, 71–84.

Boucheix, J., & Schneider, E. (2009). Static and animated presentations in learning dynamic mechanical systems. *Learning and Instruction*, *19*, 112–127.

Brunken, R., Seufert, T., & Paas, F. (2010). Measuring cognitive load. In J. L. Plass, R. Moreno & R. Brunken (Eds.), *Cognitive load theory* (pp. 181–202). New York: Cambridge University Press.

Butcher, K. R. (2014). The multimedia principle. In R. E. Mayer (Ed.). In *The Cambridge handbook of multimedia learning* (2nd ed., pp. 174–205). New York: Cambridge University Press.

Cheon, J., Crooks, S., & Chung, S. (2014). Does segmenting principle counteract modality principle in instructional animation? *British Journal of Educational Technology*, *45*, 56–64.

Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, *25*, 315–324.

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

Comenius, J. A. (1887). *Orbis pictus*. Syracuse, NY: Bardeen (Originally published in 1658.)

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, *94*, 428–434.

Crooks, S. M., Cheon, J., Inan, F., Ari, F., & Flores, R. (2012). Modality and cueing in multimedia learning: Examining cognitive and perceptual explanations for the modality effect. *Computers in Human Behavior*, *28*, 1063–1071.

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

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology*, *21*, 731–746.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from complex animation: Seeing is understanding? *Learning and Instruction*, *20*, 111–122.

Doolittle, P. E., & Altstaedter, L. L. (2009). The effect of working memory capacity on multimedia learning: Does attentional control result in improved performance? *Journal of Research in Innovative Teaching*, *2*, 7–23.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58.

Dunsworth, Q., & Atkinson, R. K. (2007). Fostering multimedia learning of science: Explaining the role of an animated agent's image. *Computers & Education*, *49*, 677–690.

Eitel, A., Scheiter, K., & Schüler, A. (2013). How inspecting a picture affects processing of text in multimedia learning. *Applied Cognitive Psychology*, *27*, 451–461.

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

Frechette, C., & Moreno, R. (2010). The roles of animated pedagogical agents' presence and nonverbal communication in multimedia learning environments. *Journal of Media Psychology*, *22*, 61–72.

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. (2013). Designing instructional text for conversational style: A meta-analysis. *Educational Psychology Review*, *25*, 445–472.

Haber, J. (2014). *MOOCs*. Cambridge, MA: MIT Press.

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

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load, and instructional animation. *Applied Cognitive Psychology*, *21*, 713–729.

Hassanabadi, H., Robatjazi, E. S., & Savoji, A. P. (2011). Cognitive consequences of segmentation and modality methods in learning from instructional animations. *Procedia – Social and Behavioral Sciences*, *30*, 1481–1487.

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

Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction*, *18*, 135–145.

Jamet, E., & Le Bohec, O. (2007). The effect of redundant text in multimedia instruction. *Contemporary Educational Psychology*, *32*, 588–598.

Jarodzka, H., van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see: Guiding students' attention via a model's eye movements fosters learning. *Learning and Instruction*, *25*, 62–70.

Jeung, H., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology*, *17*, 329–343.

Johnson, C., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, *18*, 178–191.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*, 351–371.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, *92*, 126–136.

Kartal, G. (2010). Does language matter in multimedia learning? Personalization principle revisited. *Journal of Educational Psychology*, *102*, 615–624.

Kester, L., Kirschner, P. A., & van Merrienboer, J. J. G. (2004a). Timing of information presentation in learning statistics. *Instructional Science*, *32*, 233–252.

Kester, L., Kirschner, P. A., & van Merrienboer, J. J. G. (2004b). Information presentation and troubleshooting in electrical circuits. *International Journal of Science Education*, *26*, 239–256.

Kester, L., Kirschner, P. A., & van Merrienboer, J. J. G. (2005). The management of cognitive load during complex cognitive skill acquisition by means of computer-simulated problem solving. *British Journal of Educational Psychology*, *75*, 71–85.

Kester, L., Kirschner, P. A., & van Merrienboer, J. J. G. (2006a). Just-in-time information presentation: Improving learning a troubleshooting skill. *Contemporary Educational Psychology*, *31*, 167–185.

Kester, L., Lehnen, C., van Gerven, P. W. M., & Kirschner, P. A. (2006b). Just-in-time, schematic supporting information presentation during cognitive skill acquisition. *Computers in Human Behavior*, *22*, 93–112.

Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies*, *65*, 911–930.

Kühl, T., Scheiter, K., Gerjets, P., & Edelmann, J. (2011). The influence of text modality on learning with static and dynamic visualizations. *Computers in Human Behavior*, *27* (1), 29–35.

Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology*, *25*, 943–951.

Lindow, S., Fuchs, H. M., Fürstenberg, A., Kleber, J., Schweppe, J., & Rummer, R. (2011). On the robustness of the modality effect: Attempting to replicate a basic finding. *Zeitschrift für Pädagogische Psychologie*, *25*, 231–243.

Lusk, M. M., & Atkinson, R. K. (2007). Animated pedagogical agents: Does their degree of embodiment impact learning from static or animated worked examples? *Applied Cognitive Psychology*, *21*, 747–764.

Lusk, D. L., Evans, A. D., Jeffrey, T. R., Palmer, K. R., Wikstrom, C. S., & Doolittle, P. E. (2009). Multimedia learning and individual differences: Mediating the effects of working memory capacity with segmentation. *British Journal of Educational Technology*, *40*, 636–651.

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*, *93*, 377–389.

Mayer, R. E. (2008). Multimedia literacy. In J. Coiro, M. Knobel, C. Lankshear & D. J. Leu (Eds.), *Handbook of research on new literacies* (pp. 359–376). New York: Lawrence Erlbaum Associates.

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

Mayer, R. E. (Ed.) (2014a). *The Cambridge handbook of multimedia learning* (2nd ed.). New York: Cambridge University Press.

Mayer, R. E. (2014b). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. In R. E. (Ed.). In *The Cambridge handbook of multimedia learning* (2nd ed., pp. 345–368). 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 agent. *Journal of Experimental Psychology: Applied*, *18*, 239–252.

Mayer, R. E., DeLeeuw, K. E., & Ayres, P. (2007). Creating retroactive and proactive interference in multimedia learning. *Applied Cognitive Psychology*, *21*, 795–809.

Mayer, R. E., Dow, G., & Mayer, S. (2003a). 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., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction*, *33*, 12–18.

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. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.). In *The Cambridge handbook of multimedia learning* (2nd ed., pp. 279–315). 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., & Jackson, J. (2005). The case for coherence in scientific explanations: Quantitative details hurt qualitative understanding. *Journal of Experimental Psychology: Applied*, *11*, 256–265.

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

Mayer, R. E., Mathias, A., & Wetzell, K. (2002a). 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., Mautone, P., & Prothero, W. (2002b). Pictorial aids for learning by doing in a multimedia geology simulation game. *Journal of Educational Psychology*, *94*, 171–185.

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., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles. In R. E. Mayer (Ed.). In *The Cambridge handbook of multimedia learning* (2nd ed., pp. 316–344). New York: Cambridge University Press.

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. (2003b). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology*, *95*, 419–425.

Mayrath, M. C., Nihalani, P. K., & Robinson, D. H. (2011). Varying tutorial modality and interface restriction to maximize transfer in a complex simulation environment. *Journal of Educational Psychology*, *103*, 257–268.

McLaren, B. M., DeLeeuw, K. E., & Mayer, R. E. (2011a). Polite web-based intelligent tutors: Can they improve learning in classrooms? *Computers & Education*, *56*, 574–584.

McLaren, B. M., DeLeeuw, K. E., & Mayer, R. E. (2011b). A politeness effect in learning with web-based intelligent tutors. *International Journal of Human-Computer Studies*, *69*, 70–79.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive psychology*, *41*, 49–100.

Moreno, R. (2007). Optimising learning from animations by minimising cognitive load: Cognitive and affective consequences of signaling and segmenting methods. *Applied Cognitive Psychology*, *21*, 765–781.

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

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

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. (2002a). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology*, *94*, 156–163.

Moreno, R., & Mayer, R. E. (2002b). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology*, *94*, 598–610.

Moreno, R., & Mayer, R. E. (2004). Personalized messages that promote science learning in virtual environments. *Journal of Educational Psychology*, *96*, 165–173.

Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, *19*, 309–326.

Moreno, R., Mayer, R. E., Spires, H., & Lester, J. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, *19*, 177–213.

Moreno, R., Reislein, M., & Ozogul, G. (2010). Using virtual peers to guide visual attention during learning: A test of the persona hypothesis. *Journal of Media Psychology*, *22*, 52–60.

Naumann, J., Richter, T., Flender, J., Cristmann, U., & Groeben, N. (2007). Signaling in expository hypertexts compensates for deficits in reading skill. *Journal of Educational Psychology*, *99*, 791–807.

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

Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior*, *26*, 110–117.

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

Park, B., Moreno, R., Seufert, T., & Brunken, R. (2011). Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior*, *27*, 5–10.

Pashler, H., & Wagenmakers, E.-J. (2012). Editor's introduction to the special issue on replicability in psychological science: A crisis in confidence? *Perspectives in Psychological Science*, *7*, 528–530.

Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction*, *29*, 128–140.

Pociask, F. D., & Morrison, G. R. (2008). Controlling split attention and redundancy in physical therapy instruction. *Educational Technology Research and Development*, *56*, 379–399.

Reinwein, J. (2012). Does the modality effect exist? and if so, which modality effect? *Journal of Psycholinguistic Research*, *41*, 1–32.

Rey, G. D. (2010). Reading direction and signaling in a simple computer simulation. *Computers in Human Behavior*, *26*, 1176–1182.

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

Saettler, P. (1990/2004). *The evolution of American educational technology*. Greenwich, CT: Information Age Publishing.

Saettler, P. (2004). *The evolution of American educational technology*. Greenwich, CT: Information Age Publishing (Originally published in 1990.)

Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory and Cognition*, *34*, 344–355.

Scheiter, K., & Eitel, A. (2010). The effects of signals of learning from text and diagrams: how looking at diagrams earlier and more frequently improves understanding. *In Lecture Notes in Computer Science*, *6170*, 264–270.

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010a). A closer look at split visual attention in system and self-paced instruction in multimedia learning. *Learning and Instruction*, *20*, 100–110.

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010b). Explaining the modality and contiguity effects: New insights from investigating students' viewing behaviour. *Applied Cognitive Psychology*, *24*, 226–237.

Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 72–103). New York: Cambridge University Press.

Schuler, A., Scheiter, K., & Gerjects, P. (2013). Is spoken text always better? Investigating the modality and redundancy effect with longer text presentation. *Computers in Human Behavior*, *29*, 1590–1601.

Schuler, A., Scheiter, K., Rummer, R., & Gerjets, P. (2012). Explaining the modality effect in multimedia learning: Is it due to a lack of temporal contiguity with written text and pictures? *Learning and Instruction*, *22*, 92–102.

Shavelson, R. J., & Towne, L. (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Stiller, K. D., Freitag, A., Zinnbauer, P., & Freitag, C. (2009). How pacing of multimedia instructions can influence modality effects: A case of superiority of visual texts. *Australasian Journal of Educational Technology*, *25*, 184–203.

Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*, *9*(1), 59–71.

Sung, E., & Mayer, R. E. (2012). When graphics improve liking but not learning from online lessons. *Computers in Human Behavior*, *28*, 1618–1625.

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

Tabbers, H. K., Martens, R. L., & van Merrienboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology*, *74*, 71–81.

Thorndike, E. L. (1906). *Principles of teaching based on psychology*. New York: Seiler.

Um, E., Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). Emotional design in multimedia learning. *Journal of Educational Psychology*, *104*, 485–498.

Wang, N., Johnson, W. L., Mayer, R. E., Rizzo, P., Shaw, E., & Collins, H. (2008). The politeness effect: Pedagogical agents and learning outcomes. *International Journal of Human-Computer Studies*, *66*, 98–112.

Wentzel, K. R., & Miele, D. B. (Eds.) (2016). *Handbook of motivation at school* (2nd ed.). New York: Routledge.

Witteman, M. J., & Segers, E. (2010). The modality effect tested in children in a user-paced multimedia environment. *Journal of Computer Assisted Learning*, *26*, 132–142.

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

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, *22*, 449–457.

Wouters, P., Paas, F., & Van Merriënboer, J. J. G. (2009). Observational learning from animated models: Effects of modality and reflection on transfer. *Contemporary Educational Psychology*, *34*(1), 1–8.

Yue, C. L., Bjork, E. L., & Bjork, R. A. (2013). Reducing verbal redundancy in multimedia learning: An undesired desirable difficulty? *Journal of Educational Psychology*, *105*, 266–277.