10-2019

Chapter 3: Multimedia Learning Theory and Instructional Message Design

Miguel Ramlatchan
Old Dominion University, mramlatc@odu.edu

Instructional Message Design:
Theory, Research, and Practice

Chapter 3: Multimedia Learning Theory and Instructional Message Design

Miguel Ramlatchan

Old Dominion University

Acknowledgments:

I'd like to thank my wife and kids for their patience and support, Andy and John at ODU for your encouragement and opportunity, the experienced and talented instructors I've had along the way, and thanks to our contributors, proofreaders, and reviewers!

Citation:

Ramlatchan, M. (2019). Multimedia learning theory and instructional message design. In M. Ramlatchan (Ed.), *Instructional Message Design: Theory, Research, and Practice* (Vol. 1). Norfolk, VA: Kindle Direct Publishing.

Chapter 3: Multimedia Learning Theory and Instructional Message Design

Miguel Ramlatchan

Key Points:

- Multimedia learning theory describes the use of multiple simultaneous techniques in instructional message design, such as combining narration and visuals in a presentation.
- 1) Dual coding, 2) limited working memory capacity, and 3) the need to maximize cognitive resources for learning are fundamental principles.
- The key to effective multimedia design is to minimize extraneous processing, manage essential processing, and maximizing working memory resources available for generative processing.

Abstract

Multimedia learning theory describes how the designers of instructional messages, systems, and learning environments can optimize learning. The principles and heuristics of multimedia learning theory have application in traditional and online environments, with young and adult learners, in K-12, higher education, military, corporate, government, and informal learning environments. This diversity of application is based on the foundational premise that all learners can independently process auditory and visual information, have limited working memory resources, and require cognitive resources to process new information and to learn. This chapter describes the basic tenets of multimedia learning theory, best practices that can improve our message design

and communication, and exciting future directions that we can take new research.

Introduction

When teaching students, what is better, textbooks or iPads? (iPads right?). When developing my PowerPoint slides for class, I should include a lot of color and animations and sound effects to keep my learners' attention, right? As an instructional designer, should I work to include animation or video in my project, and do those visuals require the added time and expense of narration? Designers and instructors have access to an ever increasing multitude of software functionality, online resources, and ever evolving toolsets. Though where are the research-based best practices that can guide instructional message design with these resources? Subscribing to the heuristics and principles of multimedia learning theory is one option. Multimedia learning theory provides evidence-based guidelines for creating and fostering effective communication and learning using technology. The results of nearly three decades of research can be used to help guide and inform instructors and instructional designers as they navigate the many available tools, techniques, and technologies in the search to enhance learning effectiveness.

Multimedia is the use of multiple presentation tools or techniques to deliver information. Audio and visual presentation technologies provide an effective set of tools for instructors and instructional designers to communicate with learners. Mayer's multimedia learning theory provides an informative set of principles that can be used to create effective instructional message design. It is helpful to understand the origins of multimedia learning from the original sources to also understand how to best apply the theory in practice and plan for future research. Several other theories, models, and many other research studies influenced the evolution of multimedia learning theory. However, the main contributions come from Paivio's dual coding theory, Baddeley's working memory model, and Sweller's cognitive load theory (Mayer & Moreno, 2003).

Dual Coding Theory

Paivio's dual coding theory evolved from Paivio's research on noun-adjective pairs, noun-noun pairs, and how these aspects of language appeared to evoke mental images (Paivio, 1963, 1965). In several of these early experiments, images were evoked by 'peg' words (or words intended to be used to recall other words). The general findings of these studies also suggested that concrete nouns appeared to generate related images more reliably than adjectives or abstract nouns. These vocabulary and imagery findings would evolve into Paivio's dual coding theory, which describes specialized cognitive resources used by learners to process verbal and nonverbal information (Paivio, 1969, 1971, 1986). Humans appear to have independent systems for the processing of verbal and nonverbal information. Interconnections between verbal and nonverbal information are also made and aid in knowledge recall. For instance, images can be given verbal names, and names can be associated with images. Also, single images can be associated with multiple names, and a name can be associated with multiple images (Paivio, 1991). The theory also describes what can be considered units of working memory resources called "logogens" in the verbal processing system and "imogens" in the nonverbal processing system, see Figure 1 (Clark & Paivio, 1991).

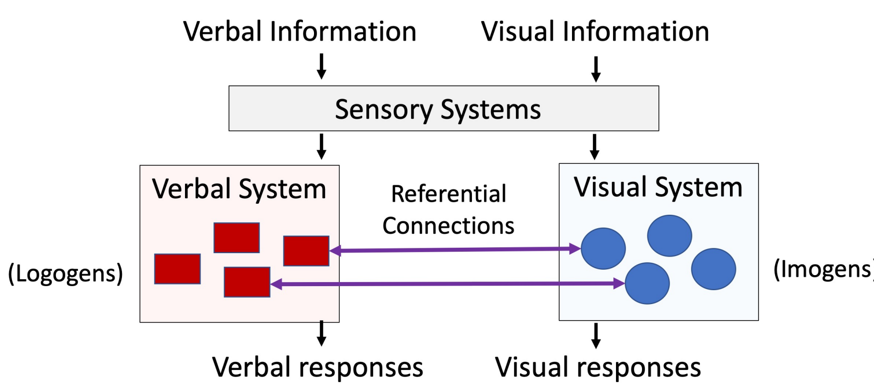

Figure 1. Paivio's dual coding theory describes logogens and imogens in verbal and visual information processing channels (modified from Paivio, 1986)

Logogens are specialized for linguistic information and imogens are specialized for nonverbal or imagery information. For instance, the spoken word "telephone" would be processed by linguistic logogens in the verbal processing system (Clark & Paivio, 1991). This processing would suggest associated imagery of telephones as well as associated sounds of telephones; this recalled nonverbal information would be processed by imogens. The two systems are able to create referential connections between logogen and imogen processed information. The result can be described as a verbal stimuli trigger to recall an entire telephone schema from long-term memory into working memory. This schema is a pattern of related ideas, words, sounds, and images that have been stored and modified over time in long-term memory. The idea that images and spoken words can be processed separately but associated together by a learner had a significant influence on multimedia learning theory (Mayer & Anderson, 1991, 1992, Mayer & Sims, 1994).

Short-term and Working Memory

Baddeley's working memory model evolved out of research into words, word length, general recall, and visual recall. It was found in a series of ten experiments that participant understanding and recall of verbally presented information was negatively affected by also having to remember six other items, but not as affected when having to recall lists of fewer than three items (Baddeley & Hitch, 1974). Baddeley & Hitch also suggested that short term memory was in actuality doing more than storing information; these cognitive resources were also being used for information processing. Thus, Baddeley and Hitch (1974) began to use the more accurate "working memory" description for cognitive resources that are apparently allocated for both short-term recall and processing. It was also found that if experiment participants rehearsed the words for themselves then they could retain those words in short term memory for an even longer length of time (as compared to not rehearsing). This result suggested a cognitive "loop." Baddeley would describe this as a phonological loop, or cognitive resources that appeared to be reserved for processing of verbal information (Baddeley, 1986).

Research into the visual aspects of working memory also began to yield similar insight into another subsystem of working memory

(Baddeley, Grant, Wright, & Thomson, 1975). It was found during this set of experiments that visual memory processing tasks did not detrimentally interfere with phonemic based recall. These early studies also suggested the potential for a "common central processor" (Baddeley and Hitch, 1974, p. 80). This central processing could be an aspect of working memory that synthesized processed information from the visual and phonologic subsystems into chunks or relationships for storage into long-term memory. Further research from these early findings continued to strongly suggest that learners could independently process both visual and phonological information and supported the existence of a central processing function (Baddeley, 1992). By the mid-90s, Baddeley's working memory model had evolved to describe two independent subsystems and central integration of these subsystems (Baddeley & Hitch, 1994). The model included a phonological loop subsystem that processes audio, a visuospatial sketchpad subsystem that processes visuals, and a central processing system for control of attention and subsystem integration.

Baddeley would specifically recall the work of Miller's seven plus or minus two units of working memory, and the use of 'chunks' to describe units of working memory (Baddeley, 1994; Miller 1956). The 'episodic buffer' aspect of central processing was later added to the model to more specifically describe the processing of visual and auditory information into chunks or 'episodes' for storage in long-term memory, see Figure 2 (Baddeley, 2000). The model that humans have limited working memory resources, used for both short term storage of information and used for actively and independently processing that information, had a substantial impact on the development of multimedia learning theory (Mayer & Moreno, 1998, 1999, 2001, 2003, Mayer, Heiser, & Lonn, 2001).

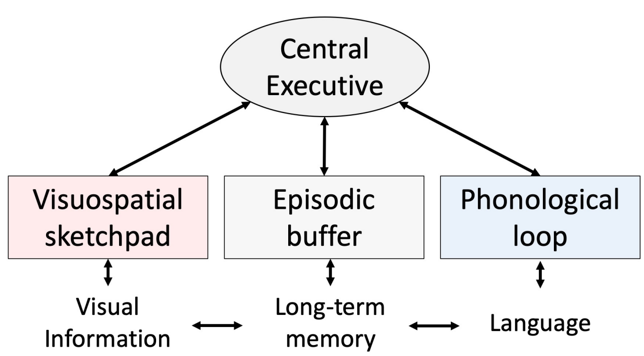

Figure 2. Baddeley's working memory model also considers the independent processing of visual and narrative information (modified from Baddeley, 2000).

Cognitive Load Theory

Sweller's cognitive load theory began with work on trigonometry word problems and the realization that students appeared less cognitively overwhelmed when they were given an example to follow during the problem-solving process (Sweller, 1988). To describe what Sweller called "cognitive processing load," Sweller notes numerous problem-solving experiments when students were more successful as the goals of the problems were simplified (Sweller, 1988, p. 263). Using a variety of physics, geometry, and maze problems, Sweller found that eliminating the implicitly stated end-goal resulted in students exploring the problem and finding the solution on their own. It appeared that not having to store problem-solving rules in working memory freed cognitive resources for working on the problems. It also appeared that the reduction of cognitive load could describe earlier experiments when learning effectiveness appeared to improve when students were given worked examples during their learning (Sweller & Cooper, 1985). Learners in these experiments did not have to store problem-solving rules in working memory (as they referred to the given example) while occupied with problem-solving. An expert has schemata stored in long-term memory that they can recall when problem-solving, novices do not and thus have to rely on inefficient "means-ends" analysis, or they focus more on the end goal (Sweller, 1989). It appeared that when students only focused on the step-by-step rules to solve the problem with only the solution as the end goal, they tended not to form the intrinsic schemata required to become experts.

Bartlett's classic experiments indicated that humans develop schema or patterns of ideas that are stored together in long-term memory as a single unit (Bartlett, 1932). It was found that when given new or unfamiliar information, such as when asked to comprehend the story the "War of the Ghosts," listeners compared the new information to their existing schemata or patterns of existing memory. British students (circa early 1930s) in this experiment did not have a schema for the Native American concept of "canoe" and so the participants translated this term as “boat” in the experiment. The unfamiliar schema was integrated into a pre-existing schema by novice learners.

Schema is a single pattern of memories that can be recalled and stored in working memory and will only occupy a single unit of working memory resources. This is analogous to Miller’s also classic description of a ‘chunk’ or unit of working memory that is also a pattern of related memories or elements also stored together as a single unit of long-term memory (Miller, 1956). Sweller uses both ‘chunks’ and ‘schema’ to describe and further an important aspect of his developing cognitive load theory, specifically that schemata storage renders human long-term storage virtually limitless (Sweller, 1994a; Mousavi, Low, & Sweller, 1995).

Sweller’s work in the early 1990s focused on what would become extraneous cognitive load, and the need for instructional designers to reduce the split attention effect and the redundancy effect (Sweller, 1991). The aspect of eliminating split attention effect would become an especially important component in what would eventually become multimedia learning theory. Split attention is the creation of extraneous cognitive load by separating relevant content in an instructional design, forcing learners to use cognitive resources to actively combine or recombine these elements in working memory. An example of reducing split attention and extraneous cognitive load would be to integrate worked examples with problems to be solved. Another classic example of the split attention effect is having a diagram on one page of a book and the text describing that diagram on another page, requiring the learner to flip back and forth between pages. This misguided instructional message design practice forces the learner to utilize cognitive resources as they flip between pages in text, thus adding extraneous cognitive load.

The term “intrinsic load” was soon added to the theory to describe the inherent difficulty of content, especially content where elements interact with each other (Sweller, 1994b). An example of high intrinsic load would be complex math problems where learners have to arrange, organize, and interact with multiple variables, and relationships between those variables, to arrive at a solution. By the late 1990s, cognitive load theory included all three of the now familiar major components of cognitive activity including extraneous load, intrinsic load, and now germane load which described the resources remaining to process relevant information (Sweller, van Merrienboer,

& Paas, 1998). This revision to cognitive load theory described a learner's working memory resources as a function and combination of extraneous, intrinsic, and germane cognitive load. For instance, an instructional designer could work to reduce split-attention effects and redundancy in instructional designs and thus reduce extraneous load. At the same time, the designer could also chunk difficult content into simpler elements in an effort to also manage intrinsic cognitive load. The result of minimizing both extraneous and intrinsic load would maximize resources for germane load, or processing of relevant information.

Sweller would continue to revise cognitive load theory, specifically revising and renaming the idea of germane cognitive "load" into germane cognitive "resources" (Sweller, Ayers, & Kalyuga, 2011, p.57). This subtle change more effectively communicates that intrinsic and extraneous processing inflicts an actual load on working memory in the form of accessible resources available for germane or relevant processing. In other words, available germane resources are a function of intrinsic and extraneous load. The theory that learners have germane resources used to process both intrinsic and extraneous information, and that a split attention effect will increase extraneous load, would be incorporated into the evolving theory of multimedia learning (Mayer, Bove, Bryman, Mars, & Tapangeo, 1996, Mayer & Moreno, 1998, 1999, Mayer, Moreno, Boire, & Vagge, 1999).

The Evolution of Multimedia Learning

Mayer's multimedia learning theory developed from research into text and illustrations and experiments that suggested that illustrations with integrated text improved learning effectiveness (Mayer, 1989). In the early 1990s, Paivio's work on dual coding theory began to inform Mayer's research with narration and animation. Mayer's results indicated that learning was most effective during treatments where the participants were able to see the animation visuals as well as hear the integrated audio narration of those visuals at the same time (Mayer & Anderson, 1991). Animation without narration and narration without animation treatments were not as effective. A further set of experiments yielded similar results when narrated animation was compared to trials of animation then narration,

narration then animation, only animation, and only narration (Meyer & Anderson, 1992; Mayer, & Sims, 1994). As dual coding describes, the learners' audio system processed the narration while the learners' visual system independently processed the animation, and central working memory resources integrated visual and narrated information into schemata. These findings were similar to the independent phonological loop and visuospatial sketchpad described by Baddeley.

Sweller and his colleagues found similar results when comparing audio integrated with visuals, as compared to the visuals alone or the audio alone (Mousavi, Low, & Sweller, 1995). Meyer integrated these findings, along with the implications of split-attention effect into another series of experiments. In a series of experimental trials, participants who viewed and listened to animation and narration outperformed participants who viewed the same animation with the text equivalent of the narration also on the screen (Mayer & Moreno, 1998). These findings were further supported by Paivio's dual coding theory and Baddeley's working memory model. Learners appeared to use dual sensory channels to process animation and available narration, though only used their visual channel when processing animation and on-screen text.

Similar findings also resulted when using different animated content, and trials with narration, integrated text, and separated text (Mayer & Moreno, 1999). This study specifically looked for results predicted by Sweller's split attention effect, or a temporal example described as a contiguity principle. The contiguity principle states that learning will be more effective when narration and visuals are timed and presented together, thus reducing or eliminating extraneous load caused by the split attention effect. The results provided further examples that narrated animation was processed more efficiently than animation with integrated text and animation with separated text.

Mayer, Baddeley, and Paivio all provide strong evidence that learners are able to process visual and audio information independently (Baddeley, 1994; Mayer & Moreno, 1999; Paivio, 1991). Mayer, Baddeley, and Sweller all provide empirical results that suggest that learners, even with independent processing, still have limited working memory resources (Baddeley, 1994; Mayer & Moreno, 1999; Sweller, van Merrienboer, & Paas, 1998). Mayer and Sweller provide evidence that presenting information with both visuals and narration can be more effective and efficient in schema creation than the same content presented with just visuals or just audio

(Mayer & Moreno, 1999; Sweller, van Merrienboer, & Paas, 1998). Taken together, these theories, experiments, and models provide the background and basis for multimedia learning theory.

Multimedia learning theory describes a series of processes that are taking place as a student is creating a new schema (Mayer, Heiser, & Lonn, 2001). The first step in the learning process is the initial viewing and listening to instructional content and the immediate storage of that information in short term memory. In this step, text is essentially visual words that when presented with diagrams then both the diagrams and the text are processed by the visual processing channel. When words are presented via audio, the narration is instead processed by the audio processing channel, while visuals are processed by the visual channel. The intrinsic content is separated from the extraneous content in the first phase of working memory. Next, the remaining germane resources in working memory create relationships between the visual and verbal information and recalls associated previous knowledge from long-term memory. Recalled schema is then compared to new information where the learner creates understanding. Finally, new schema can be created, or existing schema modified, and stored in long-term memory (see Figure 3).

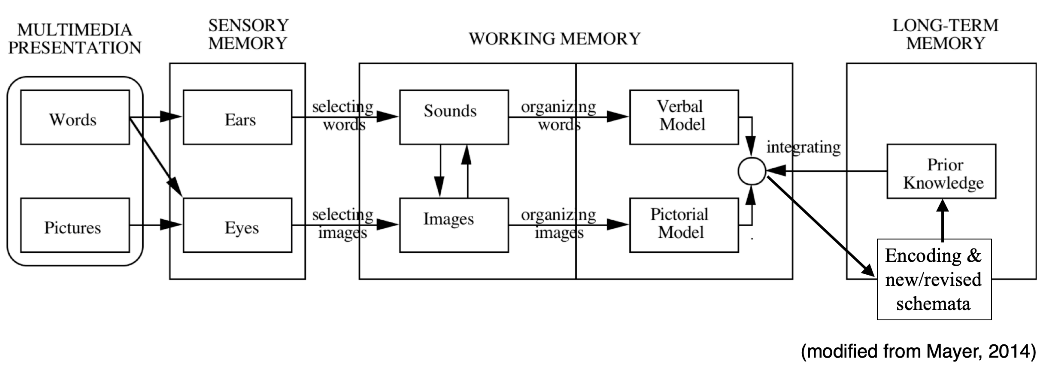

Figure 3. Multimedia learning theory describes two cognitive processing channels available to our learners, one for processing auditory information and one for processing visual information, and the result is the modification or development of new schemata in long-term memory, or learning (modified from Mayer, 2014).

By the early 2000s, Mayer's cognitive theory of multimedia learning had solidified into three main principles (Mayer & Moreno, 2003). The first principle is the assumption that learners have independent channels for verbal and visual information and using both channels simultaneously is more efficient than using either channel alone. The second principle is that the two processing channels in working memory have limited capacity for both short-term storage and active processing. The third principle is that for learning to occur, working memory must actively process, pull previous information, and create and store new or modified schema into long-term memory (see Table 1 for a summary).

Foundational Principles:

1. The **Dual Channels principle**, states that our learners have two independent cognitive systems for processing visual and auditory information,

2. The **Limited Capacity principle**, states that our learners have limited working memory resources, and

3. The **Active Processing principle**, which states that to learn students need to focus on relevant information, organize that information for themselves, and relate that information to previous schemata.

Table 1, The three foundational principles of multimedia learning theory (Clark & Mayer, 2016; Mayer & Moreno, 2003)

As with early work with new animation technology in the 1990s, Mayer continued to explore new instructional message design tools and early virtual reality applications using new multimedia learning predictions (Moreno & Mayer, 2002). Treatments using desktop monitors were compared to groups using head-mounted displays; the narrated presentations resulted in greater learning outcomes than groups viewing animations with text. These findings continued the dual coding assumptions of multimedia learning theory, and also showed that the specific technology or media used is less important than the instructional techniques and how the affordances of technology and media are used. Desktop monitors produced comparable or slightly superior results as compared to the new wearable technology, and the use of visuals and narration together were still more important in these experiments.

Media and Methodology

As in early research studies, multimedia learning theory can also apply to the use of text and diagrams (Mayer, 1989). A series of media comparison studies found that good instructional design was applicable independently of the media or the technology used to deliver that message (Mayer, 2003). Dual channel processing, limited working memory, and the need to actively create schema applies to the use of computer or paper-based message designs. In another study, it was found that when both the media and the design methodology are varied, user-controlled text with diagrams can be more effective than narrated animation without user controls (Mayer R., Hagerty, Mayer, S., Campbell, 2005). The ability for participants to review and re-review the diagrams with text was compared to treatments where participants were not able to control the playback of the narrated animation. Both the media and the design methodology were different in these experiments. However, when the media is held constant, the methodology can be adjusted to find the optimal learning effectiveness of the media.

Multimedia learning theory and the use of both audio and video can inform and predict the successful application of multimodal interactive learning environments. Results from asynchronous narrated animation or presentations should be generalizable to synchronous conferencing and online distance learning applications where audio and video is shared to and from all participants (Moreno & Mayer, 2007). The use of web conferencing would be the media being adjusted, the method of presentation is unchanged, and thus learners should benefit from the efficiency of dual coding. All things being equal, the learning effectiveness of an online synchronous presentation should be the same as an online asynchronous presentation, unless the instructor takes advantage of the real-time technology and fosters dialog and discussion with learners. Similarly, if the method remains constant, the use of different media such as comparing desktop and mobile device screens should not matter as long as students can see and hear the presentation. For instance, a specific comparison between electronic textbooks on mobile devices and traditional hardcopy, paper textbooks found no significant difference in learning effectiveness (Rockinson-Szapkiw, Courduff, Carter, & Bennett, 2013).

Multimedia learning theory provides results supporting instructional methodology being more important than instructional media. For instance, adding chapters and headings to a presentation improved learning effectiveness for both desktop and mobile device treatments groups, and both groups performed equivalently (Sung & Mayer, 2013). This study found that while students may have different preferences, learning effectiveness should not be impacted by device type but can be impacted by methodology and message design changes. Interestingly, the cultural context of instructional methodology or message also has a significant impact on the effectiveness of instructional media or technology (Sung & Mayer, 2012). The common thread through these studies is that multimedia learning theory can be successfully applied using a variety of technologies. The specific technology used to deliver an instructional message is less important than the message being communicated unless that technology allows for an affordance that the instructor can use to improve the message (Fiorella & Mayer, 2016; Mayer, 2018). For instance, consider a classroom of students with iPads. Simply reading an e-textbook on an Apple iPad should not yield any learning differences as compared to reading a physical paper and ink textbook. However, the iPad can connect to the Internet for additional learning resources. If the classroom teacher harnesses the affordances of the iPad by guiding students beyond the e-textbook to additional resources, then the iPad could improve learning effectiveness as compared to the physical textbook.

Processes, Principles, and Instructional Methods

The current iteration of multimedia learning theory advises heuristics beyond its foundational principles with three base processes and several guiding best practices. Multimedia learning theory is based in part on cognitive load theory, though while cognitive load can be described by extraneous load, intrinsic load, and germane resources, multimedia learning theory can be described by analogous cognitive processing. These processes are described as extraneous, essential, and generative processing (Clark & Mayer, 2016). Extraneous processing is the active use of cognitive resources to process and filter redundancy or distractions from multimedia designs. Essential processing is the utilization of cognitive resources that are used to process and simplify the complexity of a multimedia design. Generative processing is the process of analyzing, synthesizing, and organizing relevant information into schemata. In practice, all three forms of processing are occurring during learning. However, the goal of good instructional message design using multimedia is to minimize the resources consumed by extraneous and essential processing and to maximize the resources available for generative processing.

In addition to foundational dual channel, limited capacity, and active processing principles, an additional series of principles can be thought of as evidence-based instructional methods or design best practices (Clark & Mayer, 2016; Mayer, 2018).

To minimize extraneous processing:

1. The **Coherence principle** advises designers to avoid the use of unnecessary words, sounds, or graphics. Superfluous or irrelevant text, sound, and graphics will require unnecessary processing and use of cognitive resources.

2. The **Spatial Contiguity principle** advises designers to put text and graphics related to that text near each other in instructional message designs. The classic example of text on one page of a book and the figure being described by that text on a different page of that book causes unnecessary extraneous processing.

3. The **Temporal Contiguity principle** advocates synchronizing audio and video in presentations. Presenting audio before video or video before audio, or video and audio that are not in sync confuses and distracts learners.

4. The **Redundancy principle** states that on-screen text is distracting when audio and graphics are also used. Learners can be distracted by the redundancy of focusing and refocusing between the text and narrations when graphics are presented with text, and that text is read verbatim by a narrator. It is less distracting for a narrator not to read the on-screen text word-for-word.

5. The **Signaling principle** states that essential content can be highlighted to draw the learner's attention to it. Signaling can be used to cue learners to important content and can be highlighted text, the use of bold or italics, or visuals of an instructor pointing to specific content on a whiteboard.

To optimize essential processing:

6. The **Worked Example principle** states that a step-by-step demonstration can help reduce complexity when problem-solving. Giving students an example to follow when working through similar problems gives them guidance to refer to and focuses their essential processing.

7. The **Segmenting principle** states that a continuous complex presentation should instead be broken down into shorter more manageable chunks. Complex content can be simplified by breaking that complexity down into easier components.

8. The **Pretraining principle** suggests that key, unfamiliar terminology and definitions be given and discussed before an instructional unit. Similar to segmenting, students can be prepared for learning by first presenting them with and discussing key concepts and definitions.

9. The **Modality principle** suggests the use of audio rather than on-screen text during video, animations, or presentations. Presenting on-screen text with graphics only utilizes the visual processing capabilities of learners while using graphics with narration is more efficient as it utilizes both the learner's visual and auditory processing capabilities.

To increase resources for generative processing:

10. The **Personalization or Voice principle** advocates the use of a more conversational tone when narrating visuals as opposed to a formal, academic tone. A friendly narrative tone fosters social presence which enhances motivation for learning.

11. The **Embodiment principle** suggests the use of human-like gestures when including on-screen agents in multimedia designs. The human-like gestures and personifications enhance the perception of virtual social presence and also increases learner motivation.

12. The **Multimedia principle** suggests presenting relevant graphics with text rather than just text. Static or dynamic graphics combined with text can often communicate more effectively and efficiently than just text alone by presenting concepts and principles as a visual schema.

13. The **Engagement principle** suggests that instructors and teachers actively involve students by asking them questions during presentations. Students will learn better when actively involved in a discussion vice passively listening to a lecture.

Emerging Technologies and Applications

While multimedia learning theory was born of experiments with text and graphics, the principles can likely apply to a number of new and emerging technologies. Emerging instructional message design technologies include mobile devices, virtual reality, e-learning, online education, and digital whiteboards. Building on the philosophy of instructional methods being more important than instruction media, comparing learning on a PC workstation and learning from an Apple iPad should not make a difference. As expected, experiments with iPads have shown motivational differences over workstations, likely because learning with mobile devices means students do not have to be confined to computer labs (Sung & Mayer, 2013). However, learning effectiveness was statistically equivalent. Similar results were found in research with virtual reality headsets; the use of immersive virtual reality enhanced motivation though did not enhance learning effectiveness (Parong & Mayer, 2019). The novelty of the headsets and hand controllers could have increased motivation as compared to the more common use of PowerPoint.

E-learning and online education are now commonplace in K-12, higher education, and government, military, and corporate training. Multimedia learning theory can be used to guide and improve these learning environments through effective instructional message design (Clark & Mayer, 2016; Mayer, 2019; Sung & Mayer, 2013). These guidelines can also be used to effectively use drawings on traditional and digital whiteboards (Fiorella, Stull, Kuhlmann, & Mayer, 2018). In addition to enhancing social presence, especially in online environments, handwritten drawings appear to foster generative learning by building on the signaling and embodiment principles, or the use of human gestures to highlight content. The use of a transparent whiteboard that allows the instructor to look into the camera while drawing, enhances social presence, though does not appear to impact learning effectiveness as compared to the use of a traditional whiteboard (Stull, Fiorella, & Mayer, 2018).

Future research directions

Multimedia learning theory can be used to guide and predict the usefulness and learning effectiveness of visual and verbal presentations. It is critical that instructional message design is based on research and applied science and not fads, marketing, hype, opinion, and intuition (Mayer, 2018). As seen in previous multimedia studies, the technology or delivery media used by instructors or instructional designers is less important than what the technology conveys. As a result, paper illustrations with audio narration, animation with audio narration, static slides with narration, video with audio, or virtual reality with narration should all be effective ways to communicate and trigger efficient dual coding. The use of simultaneous verbal and visual information in a presentation is an effective communication technique regardless of the specific technology used. Thus, the principles of multimedia learning theory should be applicable to video with audio, and video with slides and audio.

Future research studies could use multimedia learning to guide the design of treatment groups in quantitative experiments that could extend the findings and applications of the theory. For instance, versions of multimedia presentations can be compared to each other to inform the use of audio and video in distance learning courses delivered online, to mobile devices. A version of an online presentation with narrated slides can be compared to a version with the instructor's video in a window with the narrated slides in a larger window on the screen, the narration and just the instructor video, and a narrated version where visuals switch between instructor video and slides. Potentially, these four treatments can be compared to a group who only listens to the narration without the visuals of the slides and a group who only has access to the slides without narration. Mayer's multimedia learning theory would predict that the narrated visual groups will perform best on comprehension post-tests, but which of the four versions will perform best? Other potential experiments could add real-time engagement with the instructor, variations of visuals of the instructor and visuals of presentation content, and study the social presence implications of longer presentations at digital and traditional whiteboards, writing tablets, and document cameras with and without a view of the instructor. These future study variations could serve to fill gaps in the multimedia knowledge base or to specifically test the potential benefits and optimal variations of integrating audio with both video and presentation content. The results of this series of studies could be used to guide and inform future instructional design techniques intended for augmented reality, virtual reality, and mobile applications.

Future multimedia studies will also benefit from new ways to measure load and processing in experiments. Self-reporting surveys and questionnaires offer an indirect means to measure load and processing. While it is possible to individually measure extraneous, intrinsic, and germane loads and resources (and thus potentially extraneous, essential, and generative processing), these measures remain indirect (Deleeuw & Mayer, 2008). The emergence, affordability, and accuracy of eye-tracking systems offer an emerging and direct means to measure cognitive load and extraneous, essential, and generative processing (Li, Wang, Mayer, & Liu, 2019; Stull, Fiorella, & Mayer, 2018; Xie, Mayer, Wang, & Zhou, 2019). In addition to potential direct measures of load and processing, eye-tracking can also inform designers on the effectiveness of signaling and the potential distractions of design decisions.

Conclusions

Multimedia learning theory builds on a number of previous theories and applies best practice heuristics that can be used to create successful instructional message design. Dual coding, working memory, and cognitive load theories, as well as early experiments comparing text and graphics, have developed into the foundation of multimedia learning theory. These foundational principles include the concept that humans have dual processing capabilities for auditory and visual information, have limited working memory resources, and require working memory resources for the processing of information and for learning. Working memory is also allocated to three cognitive processes when learning: extraneous, essential, and generative processing. Extraneous processing is the resources required to filter distractions, essential processing is required to analyze and sift through the complexity of a presentation, and remaining cognitive resources are allocated to generative processing or the creation of new schemata and learning. These multimedia learning processes are analogous to the extraneous load, intrinsic load, and germane resources described by cognitive load theory. The goal in instructional message design is to reduce the need for extraneous processing, manage essential processing, and maximize generative processing. Multimedia designs can be optimized by evidence-based best practices such as maintaining contiguity in design elements, avoiding redundancy, signaling learners, segmenting complex content, combining and using both audio and visual design elements, using a conversational tone in narrations, and engaging learners by involving them in the presentation.

The principles of multimedia learning theory can be used to enhance and improve the ways that instructional message design is used to provide learning opportunities and communication. We know that the message being conveyed to our learners by technology is more important than the technology itself. For instance, reading from a textbook should be just as effective as reading from an iPad. Only when the instructor or designer uses the affordances and advantages of the technology, do the choice and use of one technology over another become significant. Or, when the iPad users are able to take advantage of different online resources not available in the textbook, does the use of different technologies become effective. Comparing different technologies to each other when teaching the same way is futile. However, learning how different technologies can afford new and more effective ways to teach and communicate is much more beneficial and relevant. It is hard to estimate the number of instructional message designs in K-12, higher education, military, corporate, government, and informal learning environments that have benefited from the results of nearly 30 years of multimedia learning research. However, given the multitude of poor examples of design in these same environments, and the continued advance of technology, there are still many opportunities for designers to apply multimedia learning principles to help learners learn.

References

Baddeley, A. D. (1986). *Working memory*. New York, NY: Oxford University Press.

Baddeley, A. D. (1992). Working memory, *Science, 255*(5044), 556-559.

Baddeley, A. D. (1994). The magical number seven: Still magic after all these years? *Psychological Review, 101*(2), 353-356.

Baddeley, A. D., Grant, S., Wright, E., & Thomson, N. (1975). Imagery and visual working memory. In P. M. Rabbit & S. Dornic (Eds.) *Attention and performance,* (Vol. 5). New York, NY: Academic Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation, 8*, 47-89.

Baddeley, A. D., & Hitch, G. (1994). Development in the concepts of working memory. *Neuropsychology, 8*(4), 485-493.

Bartlett, F. (1932). *Remembering: A study in experimental and social psychology.* New York, NY: Cambridge University Press.

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

Clark, R. C., & Mayer, R. E. (2016). *e-learning and the science of Instruction: proven guidelines for consumers and designers of multimedia learning.* Hoboken, NJ: John Wiley & Sons, Inc.

DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology, 100*(1), 223-234.

Fiorella, L., & Mayer, R. E. (2016). Effects of observing the instructor draw diagrams on learning from multimedia messages. *Journal of Educational Psychology, 108*(4), 528-546.

Fiorella, L., Stull, A. T., Kuhlmann, S., & Mayer, R. E. (2018). Instructor presence in video lectures: The role of dynamic drawings, eye contact, and instructor visibility. *Journal of Educational Psychology.* http://dx.doi.org/10.1037/edu0000325

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.* http://dx.doi.org/10.1037/edu0000352

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research, 59*(1), 43-64.

Mayer, R. E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instructional, 12*, 125-139.

Mayer, R. E. (2014). Multimedia instruction. In J. M. Specter (Ed.), *Handbook of Research on Educational Communications and Technology* (4th ed). New York, NY: Springer Science+Business Media.

Mayer, R. E. (2018). Educational psychology's past and future contributions to the science of learning, science of instructional, and science of assessment. *Journal of Educational Psychology, 110*(2), 174-179.

Mayer, R. E. (2019). Thirty years of research on online learning, *Applied Cognitive Psychology, 33*(2), 152-159.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of dual-coding hypothesis. *Journal of Educational Psychology, 83*(4), 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*(4), 444-452.

Mayer, R. E., Bove, W., Bryman, A., Mars, & Tapangeo, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology, 88(*1), 64-73.

Mayer R. E., Hagerty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animation in multimedia instruction. *Journal of Educational Psychology: Applied, 11*(4), 256-265.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on Multimedia Learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*(1), 187-198.

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*(2), 312-320.

Mayer, R. E., & Moreno, R. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 92*(2), 358-368.

Mayer, R. E., & Moreno, R. (2001). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *The Journal of Educational Psychology, 92(*1), 117-125.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 42-52.

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*(4), 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*(3), 389-401.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81-97.

Moreno, R., & Mayer, R. E., (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology, 94*(3), 598-610.

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

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*(2), 319-334.

Paivio, A. (1963). Learning of adjective-noun paired associates as a function of adjective-noun word order and noun abstractness. *Canadian Journal of Psychology, 17*(4), 370-379.

Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired-associative learning. *Journal of Verbal Learning and Verbal Behavior, 4*, 32-38.

Paivio, A. (1971). *Imagery and Verbal Processes*. New York, NY: Holt, Rinehart, and Winston.

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

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3). 255-287.

Parang, J., & Mayer, R. E. (2019). Learning science in immersive virtual reality. *Journal of Educational Psychology, 110*(6), 785-797.

Rockinson-Szapkiw, A., Courduff, J., Carter, K., & Bennett, D. (2013). Electronic versus traditional print textbooks: A comparison study on the influence of university students' learning. *Computers & Education, 63,* 259-266.

Stull, A. T., 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. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology, 81*(4), 457-466.

Sweller, J. (1991). Evidence for cognitive load theory. *Cognition and Instruction, 8*(4), 351-362.

Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and Instruction, 4*, 295-312.

Sweller, J. (1994b). Why some material is difficult to learn. *Cognition and Instruction, 12*(3), 185-233.

Sweller, J., Ayers, P., & Kalyuga, S., (2011). *Cognitive load theory: explorations in the learning sciences, instructional systems and performance technologies*. New York, NY: Springer.

Sweller, J. & Cooper, G (1985). The use of worked examples as a substitute for problem solving in linear algebra. *Cognition and Instruction, 2*, 59-89.

Sweller, J., van Merrienboer, J. G., & Paas, F. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

Xie, H., Mayer, R. E., Wang, F., & Zhou, Z. (2019). Coordinating visual and auditory cueing in multimedia learning. *Journal of Educational Psychology, 11*(2). 235-255.