The Case for Rethinking Multimedia

Hal Hinderliter

Mayer's (2002) Cognitive Theory of Multimedia Learning (CTML) has been a mainstay of instructional design for nearly two decades, but recent developments in theoretical frameworks relevant to multimedia learning point to a reversal of many previously held assumptions (Mayer, 2005; Mayer & Fiorella, 2014; van den Broek et al., 2014). Activities that update or redefine CTML's many principles indicate the need for closer examination of the theories and principles used to support our multimedia design practices. This chapter will examine multiple issues regarding CTML that question its influential status within the instructional design community.

Instructional design for online course delivery has long taken a positive view toward the use of multimedia (Reiser & Gagné, 1983, p. 3). The term has been in use since at least 1959, with the term "media" (the plural of "medium") referencing the technological channels of distribution through which representations are made available to audiences; e.g., text, photography, audio recordings, television, streaming video. Multimedia presentations enable the simultaneous delivery of content via multiple channels, allowing learners to access a particular message in more than one way; e.g., a classroom lecture delivered as a narrated PowerPoint video, or an e-book that incorporates animated diagrams.

In Mayer's omnibus theory, multimedia is both the overarching concept as well as the key initial principle. According to Mayer's (2002) definition of a multimedia effect as stated in his cognitive theory of multimedia learning, delivering instruction as both visual and auditory content will increase learners' opportunity to absorb and retain its contents. While this idea may seem indisputable in today's media-rich environment, it represents the modern rejection of previously accepted notions that "combined audiovisual presentation is no better than auditory alone" (Penney, 1975,

H. Hinderliter (✉)
Northern Illinois University, DeKalb, IL, USA
e-mail: hal@halhinderliter.com

p. 69). This chapter reviews several theories that laid the groundwork for Mayer's treatise, then discusses recent developments and concerns over the applicability of its many principles in terms of the range of modes affected and the types of learners who may benefit. Accordingly, this chapter does not challenge Mayer's "presentation modes approach" (Mayer, 2002, p. 96) to multimedia; instead, its scope is limited to enumerating a short list of concerns regarding CTML's frequently invoked prescriptions: the design of the experiments that form the basis of the theory, the relevance of the situations in which these experiments were conducted, the difficulty of replicating these results, the cognitivist assumptions on which its principles are based, and the impact of these assumptions on the accessibility of CTML-guided content.

Review

The study of multimedia's efficacy for instructional tasks predates the Internet, online learning, and the use of personal computers in classrooms. Throughout Mayer's most influential research (see Mayer et al., 1996; Mayer, 2002, 2005; Mayer & Johnson, 2008), multiple aspects of multimedia learning are seen within a cognitivist framework. This focus on cognition when examining learning processes, that is, how our brains process then store information, can be traced to Miller (1956), among others. His seminal paper on the limits of short-term working memory put forth the idea that humans face "severe limitations on the amount of information that we are able to receive, process, and remember" before postulating that these limitations could be reduced by "organizing the stimulus input simultaneously into several dimensions and successively into a sequence of chunks" (Miller, 1956, p. 96). By foregrounding the importance of human cognitive processes, Miller advanced the idea that human capacity to learn was limited by inadequate working memory.

The Roots of CTML

Miller's ideas were expanded upon through Atkinson and Shiffrin's (1968) information processing theory, Anderson's (1977) schema theory, and others (e.g., Fleming & Levie, 1978; Johnson-Laird, 1983). Among the most influential of these ideas was Paivio's (1986) dual-coding theory, which proposed separate processing centers for language and nonverbal stimuli. Paivio theorized that presentation modalities, which he defined as verbal or nonverbal, have an important impact on students. Dual-coding theory differed from prior behaviorist notions of knowledge acquisition, which held that modality was unimportant; prior to the advent of cognitivism, the content being communicated was considered key, while the choice of delivery method was thought to be inconsequential.

Sweller (1988) described a production model for human problem solving that views the human mind's cognitive processes as a series of switch gates on a circuit board; this theory of cognitive load proposed a discrete, limited capacity for human visual and auditory input. Three forms of cognitive processing are summed to define learners' total cognitive load: intrinsic load, that is, the effort required to understand the primary learning task; extraneous load, that is, the undesirable additional stress incurred by poorly formed instructions; and germane load, for example, the effort involved in fixing new information within long-term memory. Cognitive load theory cautions against overloading the brain's processing capacity so that problem-solving schemas may be acquired for transfer into long-term memory. The resilience of cognitive load theory can be seen today when instructional designers configure learning environments so as to avoid overwhelming learners' restricted audiovisual processing capacity, a situation known as *cognitive overload*. The popularity of cognitive load theory led to the proposal of several related ideas, each focused on avoiding or controlling the effects of cognitive load.

Chandler and Sweller (1991) proposed the redundancy principle, that is, that redundant material decreases the intelligibility of instruction by overloading learners' processing capacity. As technology made it easier to combine text and graphics during instruction, Chandler and Sweller (1992) also identified a split attention effect as the cognitive load created by switching between focal points. By the mid-1990s, cognitive load-related theories pertaining to how learners acquire information from presentations were multiple. These closely related ideas were often used interchangeably to justify a common belief: that displaying too much text within a narrated presentation would overload visual working memory, leading to inferior learning outcomes.

Defining the Principles of Multimedia Learning

With so many related ideas competing for attention, it is unsurprising that Mayer decided to coordinate them all beneath a single umbrella, which he called the Cognitive Theory of Multimedia Learning (CTML). Mayer developed CTML over a series of publications. The first of these (Mayer et al., 1996) promoted the combining of visual and verbal stimuli in an annotated illustration called a multimedia summary. This technique was said to promote the retention and transfer of scientific information by reducing students' cognitive load through three principles: *conciseness*, meaning that only a few sentences and illustrations are used; *coherence*, meaning that related content is presented in cause-and-effect sequence; and *coordination*, meaning that graphics and sentences are presented contiguously.

In a literature review, Mayer (2002) elaborated his first list of nine key principles for multimedia learning; these principles were retained when Mayer (2014) later expanded the list to 12 principles:

- The *multimedia effect* is the belief in better transfer when a message contains words and pictures rather than words alone.
- The *spatial contiguity effect* is the belief that words and related graphics should be presented in close physical proximity to each other.
- The *temporal contiguity effect* states that visual content and related audio content should be presented at the same time.
- The *coherence effect* is that irrelevant words, graphics, and sounds should be excluded.
- The *modality effect* is the belief that animated graphics and videos should be accompanied only by narration, not by text.
- The *redundancy effect* states that animated graphics and videos should be accompanied only by narration, not the combination of text and narration.
- The *pretraining effect* says that introducing topics before instruction is better than explaining them after instruction.
- The *signaling effect* promotes signaling the importance or relationship of concepts, either graphically, for example, through the use of bullet points, or verbally, for example, using words such as "because" or "as a result."
- The *personalization principle* calls for the use of conversational language as opposed to formal language.

Expanding the Definition of CTML

In one chapter of a book he edited for Cambridge University Press, Mayer (2005) published a more detailed explanation of multimedia learning that included three primary assumptions: *dual channels*, that is, that visual and auditory stimuli are processed separately; *limited capacity*, that is, that the amount of information these channels can process is circumscribed; and *active processing*, that is, that attending to, organizing, and integrating information leads to meaningful learning.

More crucially to the purpose of this chapter, Mayer (2005) also included research that showed measurable benefits from specific forms of redundancy under certain conditions, for example, when text-only slides are accompanied by redundant narration. Mayer's acknowledgment of redundant narration's value in text-only situations advanced his prior research with Moreno (Moreno & Mayer, 2002), which reported a threefold increase in correct answers when learners were shown presentations where the narration was an exact reflection of the on-screen text, i.e., verbally redundant. This finding ran counter to Kalyuga et al.'s (1999) delineation of a redundancy effect as well as Mayer's (2002) initial definition of CTML's redundancy principle.

Soon after, Mayer and Johnson (2008) offered a more articulated sense of multimedia learning principles in an attempt to explain such contradictory test results. Mayer called these exceptions to the rule *boundary conditions*. As one occurrence of such a condition, Mayer and Johnson (2008) said that redundancy, previously defined as a deleterious effect, becomes helpful in narrated presentations when short

text labels are adjacent to the graphics they describe. They described such boundary conditions as a "reverse redundancy effect" that can occur under several conditions: when the narration is complex or contains unfamiliar words, when the narration is not in the learner's native language, when the pace of presentation is slow or learner-controlled, or when the audience is composed of low-knowledge learners. Later, Mayer and Fiorella (2014) reiterated redundancy's boundary condition for text-only slides accompanied by matching narration, noting: "The redundancy effect can disappear when no graphics are presented. In this case, adding on-screen text does not create split attention because there is no other material to process in the visual channel" (p. 299).

Other researchers have proposed additional boundary conditions, including for the personalization effect. For example, results for higher-knowledge learners did not improve when formal language was replaced by conversational language (McLaren et al., 2011; Wang et al., 2008). As CTML expanded from 9 (Mayer, 2002) to 12 principles (Mayer, 2017), each with the possibility of boundary conditions that might cause a reversal of the stated effect, its complexity has become a matter of concern. de Jong (2010) challenged the legitimacy of multimedia learning theory due to the growing number of proposed boundary conditions. While each new extension to CTML demands attention, this added complexity also increases the need for assurance of its utility. To that end, numerous questions can and should be posed to help us rethink Mayer's broadly accepted guidelines for the use of multimedia.

Discussion

Having listed some of the formative research and publications that define CTML, this chapter will now address several concerns regarding this influential theory.

Experiment Design

Unlike some theories that explore a solitary phenomenon through a single experiment, Mayer's CTML is a coalition of multiple theories that pre-date its introduction. In that 2002 publication, Mayer supported his list of principles by referring to sixty previous tests conducted within 20 studies, a corpus that he had collected and analyzed over the course of 12 years. Mayer is the principal or second investigator in nearly all of these studies, the great majority of which were conducted between 1989 and 2001. Sixty is certainly an impressive number of tests, but many of these experiments are based on the same instrument even when the stated purpose of the experiment differs from the principle it is used to support.

As just one example, multiple aspects of CTML – the multimedia effect, temporal contiguity, and pretraining – are explored through various repetitions of the

"pumps" experiment (Mayer & Gallini, 1990; Mayer & Anderson, 1991, 1992; Mayer & Mathias, 2001). In Mayer and Anderson's (1991) experiment, only 15 subjects participated in each of two treatments; in both forms of the experiment, subjects viewed a computer-based animated cut-away line drawing of the inner workings of a bicycle pump, with one group hearing an audio explanation of the process before viewing the animation, and another hearing the audio while the animation was playing. No visual text was provided. After the treatment, participants were assessed via four open-ended questions posing hypothetical situations intended to test transfer of the knowledge gained by watching the animation. Each participant was given only 2.5 min to answer each question; these responses were scored on a scale of 1–4 for the first three questions but only a maximum of two points was allocated for the fourth question. Given the brief window of time allowed for participants to compose their written responses and the subjectivity of scoring them, the truncated scale on which the responses were scored, a pretreatment screening process that did not inquire as to participants' knowledge of fluid mechanics, the unknown variations in English proficiency and compositional ability among participants, and the limited number of participants undergoing each treatment ($n = 15$), questions may be raised as to whether these results are truly representative of the phenomenon under study.

Perhaps more significantly, Mayer and Anderson (1991) state that this experiment is designed to evaluate Paivio's dual-coding hypothesis against two alternate suppositions: the single-code hypothesis, and the separate dual-code hypothesis; these alternatives to Paivio's theory are not credited to any prior source, implying that they may have been crafted to function as straw men in this scenario. These issues would already pose significant concern when considering this intended application of the experiment, but Mayer (2002) relies on the Mayer and Anderson (1991) results to support his formulation of only marginally related principles, for example, temporal contiguity. When today's instructional designers defer to CTML's temporal contiguity principle as defined by Mayer, that is, that best results come when "corresponding words and pictures are presented at the same time" (Mayer, 2002, p. 111), most are likely unaware that this position is based on treatments using audio narration with an animation, not the use of visual text with narration as commonly found in narrated text-only presentations.

Another aspect of Mayer and Anderson's (1991) experiment that fails to receive sufficient scrutiny is its reliance on system-driven timing. The efficacy of instruction via online and assistive technologies is affected by the degree of user control allowed over the speed and timing of the presentation; this is especially pertinent for non-native language speakers as well as learners with physical or neurological limitations. Other early computer-mediated experiments referenced by Mayer (2002) would have generated higher cognitive loads through the use of system-paced presentations, for example, when Mayer and Anderson's (1991) participants were given only 30–45 s to absorb the meaning of each animation. Unfortunately, Mayer (2002) does not discuss the impact of (the now ubiquitous) user-controlled timing on the validity or applicability of CTML's principles.

Situational Relevance

To say that much has changed about the delivery of multimedia instructional content since the last decade of the twentieth century is a substantial understatement. The 60 experiments offered in support of CTML include paper-based treatments (Mayer, 1989; Mayer & Gallini, 1990; Mayer et al., 1996) as well as computer-based treatments involving HyperCard stacks on a monochrome Macintosh Ilci computer (Mayer & Anderson, 1991). Additionally, these experiments took place in classrooms and computer labs rather than in the solitary isolation experienced by today's online learners, who typically access their instruction from home via the Internet. Perhaps most relevantly, students watching computer-based animations in the early 1990s would have been fascinated by such a novel high-tech approach; in today's media-rich environment, however, students are highly acclimated to the use of animation, video, and narration. Any assumption that the differences between today's educational environment and the situations in which these experiments were conducted more than two decades ago should have no influence on the validity of CTML's principles seems unlikely to be true.

To this end, Tabbers et al. (2004) asked if broadly accepted findings regarding the modality effect might not be generalizable due to the unique conditions and content of the previous research. Tabbers et al. surveyed the previous research in this area, then questioned if these experiments were adequately reflective of real educational environments (p. 74). The authors expressed concern that many of the landmark studies in this area were conducted under laboratory conditions and involved only brief instruction focused solely on technical domains (e.g., Jeung et al., 1997; Kalyuga et al., 1999; Mayer & Moreno, 1998; Moreno & Mayer, 1999; Mousavi et al., 1995; Tindall-Ford et al., 1997). This potential inapplicability of previous research to today's online educational settings exacerbates the lack of targeted research into multimedia learning's potential. Since the publication of Tabbers et al.'s provocative research, the use of computer-based experiments has grown but few previous studies have been thoroughly replicated using adequate sample sizes within modern online instructional environments.

Replication Concerns

The landmark results of early research into cognitive learning principles have proven difficult to replicate, inspiring questions as to whether we should strictly adhere to CTML's principles when multiple empirical studies can offer only a mixed record of support. As early as 1975, Penny reviewed published studies showing that learners could best remember lists in short-term memory when the information was presented auditorily rather than visually, that is, the modality effect – but contemporaneously, dissenting studies found superior results from visual presentation (Kroll et al., 1972; Marcer, 1967; Scarborough, 1972). Since that time, numerous

experiments designed to investigate individual multimedia effects have produced contrary or inconclusive results (Jeung et al., 1997; Kalyuga et al., 1999; Leahy & Sweller, 2011; Savoji et al., 2011; Tabbers et al., 2004) that cast doubt on the immutability of CTML's oft-cited principles. Research conducted by Tabbers et al. (2004) found that use of visual learning material was superior to audio in terms of student transfer and retention; in Tabber's words, "Replacing visual text with spoken text even had a negative effect on learning, contrary to what both cognitive load theory and Mayer's theory of multimedia learning would predict" (p. 80). Such challenges with replicability should be of great concern, especially given that Mayer has associated CTML compliance with large effect sizes (e.g., Mayer, 2002; Moreno & Mayer, 1999.)

Cognitivist Assumptions

Constructivist scholars have long been uneasy with the presumption that cognitive research should be seen as a deterministic force in education (Greeno, 1989; Derry, 1992). Constructivist pedagogy recommends the customization of lesson plans in order to suit the unique individuality of each learner, while cognitivism seeks to determine a singular "true" manner in which students learn new information en route to the development of global prescriptions that will benefit all. Constructivists may consider the conceptualization of students' minds as analogous to computers to be a gross oversimplification, but CTML is among many theories advanced by cognitivist thinkers that remain embedded within modern instructional design practices. This has occurred because such principles are truly useful tools in guiding the development of instructional content; however, cognitivist assumptions may be better thought of as rough outlines marking the complex contours of human understanding. This is especially true when researchers are tempted to forego experimental research by relying solely on insights gleaned from models of cognitive processes, as the complexity of real-world situations and the immense variability of human functionality cannot be accounted for through one-size-fits-all suppositions.

Impact on Accessibility

Due to the factors previously discussed, the broad applicability of CTML's principles is often in conflict with issues of accessibility. Mayer and Johnson (2008) proposed that boundary conditions are relevant to low-knowledge learners, but what of learners with other challenges, for example, vision impairment, hearing impairment, cognitive impairments, or low language proficiency? While such groups have traditionally been overlooked in research on multimedia learning, together they comprise a substantial, growing portion of the student body. The narrow focus of most cognitivist-derived educational research on neurotypical native English

speakers raises concerns regarding the generalizability of such findings. In an experiment with non-native English speakers, Toh et al. (2010) found that learners exposed to temporally contiguous and verbally redundant instruction performed significantly better than those who experienced only audio narration. Surprisingly, some researchers have decried attempts to elevate the needs of second-language students or learners with disabilities. For example, the influential scholar Sweller (2005) has strongly advocated against the use of fully redundant text and narration as a waste of precious cognitive resources; despite admitting that "information that is redundant for one person may be essential for another" (Sweller, 2005, p. 165), he remains adamant that "information should be presented in a single form only, i.e., with all other versions and all unnecessary explanation eliminated" (p. 167).

This issue foregrounds the need for accessibility in online instruction. Relevant teaching theories such as Universal Design for Learning (Rose et al., 2006) urge us to adopt the most broadly accessible approach in every situation, as opposed to the most familiar, the most convenient, or the most exclusive. Many aspects of CTML are compatible with accessible learning frameworks such as Universal Design for Learning. The multimedia effect, contiguity, coherence, pretraining, signaling, and personalization are all helpful to learners with disabilities; only CTML's modality and redundancy effects preclude the use of narration with redundant on-screen text. Regardless of its impact on neurotypical students' test scores, the use of verbally redundant multimedia presentations frees hearing-impaired learners, students with low-language proficiency, and those studying in noisy environments from the need for captions. (In such cases, however, closed captions should still be made available for use with assistive technologies such as Braille terminals.)

Conclusion

Ongoing research that explores delivery styles for multimedia presentations should be considered fundamental to our practices. Advances in computer processing power, learning management systems, interactive programming technologies, and students' familiarity with multimedia render experiments from more than a decade ago unsuitable as proxies for the modern distance learning experience. Mayer's Cognitive Theory of Multimedia Learning has provided a useful framework for the development of online instruction. However, it remains a theory in flux, with multiple studies providing contrasting insights into the utility and efficacy of its many principles. Rather than adopt an unquestioning allegiance to CTML, today's instructional designers should evaluate each principle's effectiveness on a broad variety of learners within the context of modern online learning environments. Content creators must decide whether to design for an idealized audience with the caveat that boundary conditions may apply to others, or to consider just how few learners match this idealized conception – then design presentations that offer accessibility for the many.

To quantify the validity and applicability of Mayer's multimedia learning principles to both general and nontraditional audiences, continued research is needed. Specifically, more quantitative research must be conducted on a variety of learners within actual online courses so that our accepted approaches to multimedia design can be contrasted with recent scholarship expanding or challenging those practices. Further discussion regarding the implication of multimedia theory on current practices should provide meaningful insights into our fundamental assumptions regarding instructional design and their impact, if any, on learners' underlying cognitive processes.

References

Anderson, R. C. (1977). The notion of schemata and the educational enterprise. In R. C. Anderson, R. J. Spiro, & W. E. Montauge (Eds.), *Schooling and the acquisition of knowledge*. Erlbaum.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89–195). Academic Press.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293–332.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology, 62*, 233–246.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science, 38*(2), 105–134.

Derry, S. J. (1992). Beyond symbolic processing: Expanding horizons for educational psychology. *Journal of Educational Psychology, 84*(4), 413–418.

Fleming, M., & Levie, W. H. (1978). *Instructional message design: Principles from the behavioral sciences*. Educational Technology Publications.

Greeno, J. G. (1989). Situations, mental models, and generative knowledge. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 285–318). Erlbaum.

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

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press.

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

Kroll, N. E. A., Parkinson, S. R., & Parks, T. E. (1972). Sensory and active storage of compound visual and auditory stimuli. *Journal of Experimental Psychology, 95*, 32–38.

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

Marcer, D. (1967). The effect of presentation method on short-term recall of CCC trigrams. *Psychonomic Science, 8*, 335–336.

Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology, 81*, 240–246.

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation, 41*, 85–139.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 201–212). Cambridge University Press.

Mayer, R. E. (2014). Multimedia instruction. In J. Michael Spector, M. David Merrill, E. Jan, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 385–399). Springer.

Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning, 33*, 403–423.

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., Bove, W., Bryman, A., Mars, R., & Tapangco, 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., & 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.), *The Cambridge handbook of multimedia learning* (pp. 279–315). Cambridge University Press.

Mayer, R. E., & Gallini, J. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*, 715–726.

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology, 100*(2), 380–386.

Mayer, R. E., & Mathias, A. (2001, April). *Static and interactive pretraining that reduces cognitive load during multimedia learning.* Paper presented at the annual convention of the American Educational Research Association, Seattle, WA.

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.

McLaren, B. M., DeLeeuw, K. E., & Mayer, R. E. (2011). A politeness effect in learning with web-based intelligent tutors. *International Journal of Human-Computer Studies, 69*(1–2), 70–79.

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

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91*(2), 358–368.

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94*(1), 156–163.

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

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

Penney, C. G. (1975). Modality effects in short-term verbal memory. *Psychological Bulletin, 82*(1), 68–84.

Reiser, R. A., & Gagné, R. M. (1983). *Selecting media for instruction*. Educational Technology.

Rose, D., Harbour, W., Johnston, C. S., & Abarbanell, L. (2006). Universal design for learning in postsecondary education: Reflections on principles and their application. *Journal of Postsecondary Education and Disability, 19*(2), 135–151.

Savoji, A. P., Hassanabadi, H., & Fasihipour, Z. (2011). The modality effect in learner-paced multimedia learning. *Procedia – Social and Behavioral Sciences, 30*, 1488–1493.

Scarborough, D. L. (1972). Stimulus modality effects on forgetting in short-term memory. *Journal of Experimental Psychology, 95*, 285–289.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257–285.

Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159–168). Cambridge University Press.

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

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied, 3*(4), 257–287.

Toh, S. C., Munassar, W. A. S., & Yahaya, W. A. J. W. (2010). Redundancy effect in multimedia learning: A closer look. In *Proceedings of ASCILITE – Australian Society for Computers in Learning in Tertiary Education Annual Conference 2010*, pp. 988–998.

van den Broek, G. S. E., Segers, E., & Verhoeven, L. (2014). Effects of text modality in multimedia presentations on written and oral performance. *Journal of Computer Assisted Learning, 30*(5), 438–449.

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.