The Cambridge Handbook of Multimedia Learning (3rd ed.)

2 authors:

1 Introduction to Multimedia Learning*

RICHARD E. MAYER AND LOGAN FIORELLA

INTRODUCTION

People can learn more deeply from words and pictures than from words alone. This seemingly simple proposition – which we refer to as the multimedia learning hypothesis – is the motivating idea for *The Cambridge Handbook of Multimedia Learning: Third Edition*.[1] Each of the 46 chapters in this *Handbook* examines an aspect of the multimedia learning hypothesis. In particular, multimedia researchers are interested in how people learn from words and pictures, and in how to design multimedia learning environments that promote learning. In this chapter, we provide a definition of multimedia learning, offer a rationale for multimedia learning, outline the research base for multimedia learning, summarize emerging trends reflected in this new edition, and draw distinctions between two approaches to multimedia design, three metaphors of multimedia learning, three kinds of multimedia learning outcomes, and two kinds of active learning.

Although you may think of multimedia instruction as a product of the digital age, multimedia instruction has a long history dating back to 1657, when John Comenius produced the world's first children's picture book, *Orbis Pictus* (which means "The World in Pictures" or "The World Illustrated"). Each page contained a black and white line drawing of an aspect of the world ranging from the parts of a house, to elements in a barber shop, to the types of water birds, to the parts of a plant. Each element was numbered, and a corresponding legend gave its name and description in Latin and in the child's native language, such as exemplified in Figure 1.1. As noted in the preface to an English-language version (Comenius,

1887), the book became the most popular textbook in Europe for a century. As you can see, the world's first illustrated schoolbook is based on the multimedia principle, as articulated in the preface (p. vii): "the teaching of words and things must go together." There was no research base to guide Comenius's efforts hundreds of years ago, but as you will see in this *Handbook*, a dedicated corps of multimedia researchers is building evidence-based principles to guide anyone interested in designing effective multimedia instruction in the twenty-first century. This *Handbook* is intended both to give you insights into the current state of research on multimedia learning and to help guide future work.

WHAT IS MULTIMEDIA LEARNING?

Let's begin by defining some key terms. Table 1.1 summarizes definitions of multimedia, multimedia learning, and multimedia instruction.

Multimedia

The term *multimedia* conjures up a variety of meanings. You might think of watching a YouTube video on your smartphone or playing a strategy game on your tablet – that is, multimedia as a handheld experience. You might think of sitting in a room where images are presented on one or more screens and music or other sounds are presented via speakers – that is, multimedia as a "live" performance. Alternatively, you might think of sitting at a computer screen that presents graphics on the screen along with spoken words from the computer's speakers – that is, multimedia as an online lesson. Other possibilities include watching a video on a TV screen while listening to the corresponding words, music, and sounds; or even putting on a head-mounted display and exploring an immersive environment in virtual reality in which characters talk to you. Low-tech examples of multimedia include a *chalk-and-talk* presentation where a speaker draws or writes on a blackboard (or uses an overhead projector) while presenting a lecture or a textbook lesson consisting

[1] There may be some conditions in which words or pictures alone are better than words and pictures combined, such as the redundancy effect described by Kalyuga and Sweller in Chapter 16 or the expertise reversal effect described by Kalyuga in Chapter 13.

The Air.   VI.   Aër.

A cool *Air*, 1.
breatheth gently.
  The *Wind*, 2.
bloweth strongly.
  A *Storm*, 3.
throweth down Trees.
  A *Whirl-wind*, 4.
turneth it self in a round
compass.
  A Wind *under Ground*, 5.
causeth an *Earthquake*.
  An Earthquake causeth
gapings of the Earth,
(and falls of Houses.) 6.

Aura, 1.
spirat leniter.
  Ventus, 2.
flat valide.
  Procella, 3.
sternit Arbores.
  Turbo, 4.
agit se in gyrum.

  Ventus *subterraneus*, 5.
excitat *Terræ motum*.
  Terræ motus facit
Labes (& ruinas.) 6.

Figure 1.1 Example page from Comenius' *Optis Pictus*

Table 1.1 Definitions

Term	Definition
multimedia	presenting words (such as printed text or spoken text) and pictures (such as illustrations, photos, animation, or video)
multimedia learning	building mental representations from words and pictures
multimedia instruction	presenting words and pictures that are intended to promote learning

of printed text and illustrations. In sum, most academic learning situations involve multimedia presentations because students encounter words and graphics.

We define multimedia (or multimedia message) as presenting both words and pictures. By words, we mean that the material is presented in *verbal form*, such as using printed text or spoken text. By pictures, we mean that the material is presented in *pictorial form*, such as using static graphics, including illustrations, graphs, diagrams, maps, or photos, or using dynamic graphics, including animation, video, or immersive virtual reality. This definition is broad enough to include all the scenarios described in the previous paragraph – ranging from multimedia encyclopedias to online educational games to textbooks.

If multimedia (or multimedia message) involves presenting material in two or more forms, then an important issue concerns how to characterize a form of presentation. Three solutions to this problem are the delivery media view, the presentation modes view, and the sensory modalities view. According to the delivery media view, multimedia requires two or more delivery devices, such as computer screen and amplified speakers or a projector and a lecturer's voice. According to the presentation modes view, multimedia requires verbal and pictorial representations, such as on-screen text and animation or printed text and illustrations. According to the sensory modalities view, multimedia requires auditory and visual senses, such as narration and animation or lecture and slides.

We reject the delivery media view because it focuses on the technology rather than on the learner. Instead, we opt for the presentation modes view, and to some extent the sensory modalities view. The presentation modes view allows for a clear definition of multimedia – presenting material in verbal and pictorial form – and is commonly used by multimedia researchers (Mayer, 2021). The presentation modes view is also the basis for Paivio's (1986, 2007) dual-coding theory as well as theories of multimedia learning presented in this *Handbook* (Chapter 5 by Mayer; Chapter 6 by Paas and Sweller; Chapter 7 by Schnotz; and Chapter 8 by Kester and van Merriënboer). The sensory

modalities view is also relevant because words can be presented as printed text (initially processed visually) or as spoken text (initially processed auditorily), whereas pictures are processed visually. In conclusion, as shown in Table 1.1, multimedia (or multimedia message) refers to communicating through words and pictures.

Multimedia Learning

Multimedia learning occurs when people build mental representations from words (such as spoken text or printed text) and pictures (such as illustrations, photos, animation, or video). As you can see in these definitions, multimedia refers to the presentation of words and pictures, whereas multimedia learning refers to the learner's construction of knowledge from words and pictures. The process by which people build mental representations from words and pictures is the focus of Mayer's cognitive theory of multimedia learning (Mayer, 2021; see also Chapter 5), Sweller's cognitive load theory (Sweller, Ayres, & Kalyuga, 2011; see also Chapter 6), Schnotz's integrative model of text and picture comprehension (Schnotz & Bannert, 2003; see also Chapter 7); and, to some extent, van Merriënboer's four-component instructional design theory (van Merrienboer & Kirschner, 2013; see also Chapter 8).

Multimedia Instruction

Multimedia instruction (or multimedia instructional message) involves presenting words and pictures that are intended to promote learning. In short, multimedia instruction (or multimedia instructional message) refers to designing multimedia learning environments in ways that help people build mental representations that support performance on subsequent tasks. The 30 instructional design principles described in Parts III–VII of the *Handbook* suggest ways of creating multimedia lessons intended to promote multimedia learning; and in Part VIII we find examples of how the principles can be applied in a variety of media contexts ranging from instructional video to computer games to virtual reality to online pedagogical agents.

WHAT IS THE RATIONALE FOR MULTIMEDIA LEARNING?

What is the value of adding pictures to words? Do students learn more deeply from words and pictures than from words alone? Do students engage in different learning processes when they receive words and pictures rather than words alone? These questions are essential to the study of multimedia learning. For example, suppose we asked you to listen to a short explanation of how a bicycle tire pump works: "When the handle is pulled up, the piston moves up, the inlet valve opens, the outlet valve

closes, and air enters the lower part of the cylinder. When the handle is pushed down, the piston moves down, the inlet valve closes, the outlet valve opens, and air moves out through the hose." Then, we ask you to write down an explanation of how a bicycle tire pump works (i.e., retention test) and to write answers to problem-solving questions such as "Suppose you push down and pull up the handle of a pump several times but no air comes out. What could have gone wrong?" (i.e., transfer test). If you are like most of the students in our research studies (Mayer & Anderson, 1991, 1992), you remembered some of the words in the presentation (i.e., you did moderately well on retention) but you had difficulty in using the material to answer problem-solving questions (i.e., you did poorly on transfer).

In contrast, suppose we showed you only an animation of a bicycle tire pump that depicts the actions in the pump as the handle is pulled up and then as the handle is pushed down. Frames from the animation are shown in Figure 1.2. If you are like most students in our research studies (Mayer & Anderson, 1991, 1992), you would not do well on a retention test or on a transfer test.

Finally, consider the narrated animation summarized in Figure 1.3. In this situation, you hear the steps described in words and see the steps depicted in the animation. When words and pictures are presented together as in a narrated animation, students perform well both on retention and transfer tests (Mayer & Anderson, 1991, 1992). In particular, when we focus on tests of problem-solving transfer – which are designed to measure the student's understanding of the presented material – students perform much better with words and pictures than from words alone. This pattern was found in 13 out of 13 studies, yielding a median effect size of d = 1.35 (Mayer, 2021). We refer to this finding as the *multimedia principle*, and it is examined in detail in Chapter 11.

The multimedia principle epitomizes the rationale for studying multimedia learning. There is reason to believe that – under certain circumstances – people learn more deeply from words and pictures than from words alone. For hundreds of years, the major format for instruction has been words – including lectures and books. In general, verbal modes of presentation have dominated the way we convey ideas to one another and verbal learning has
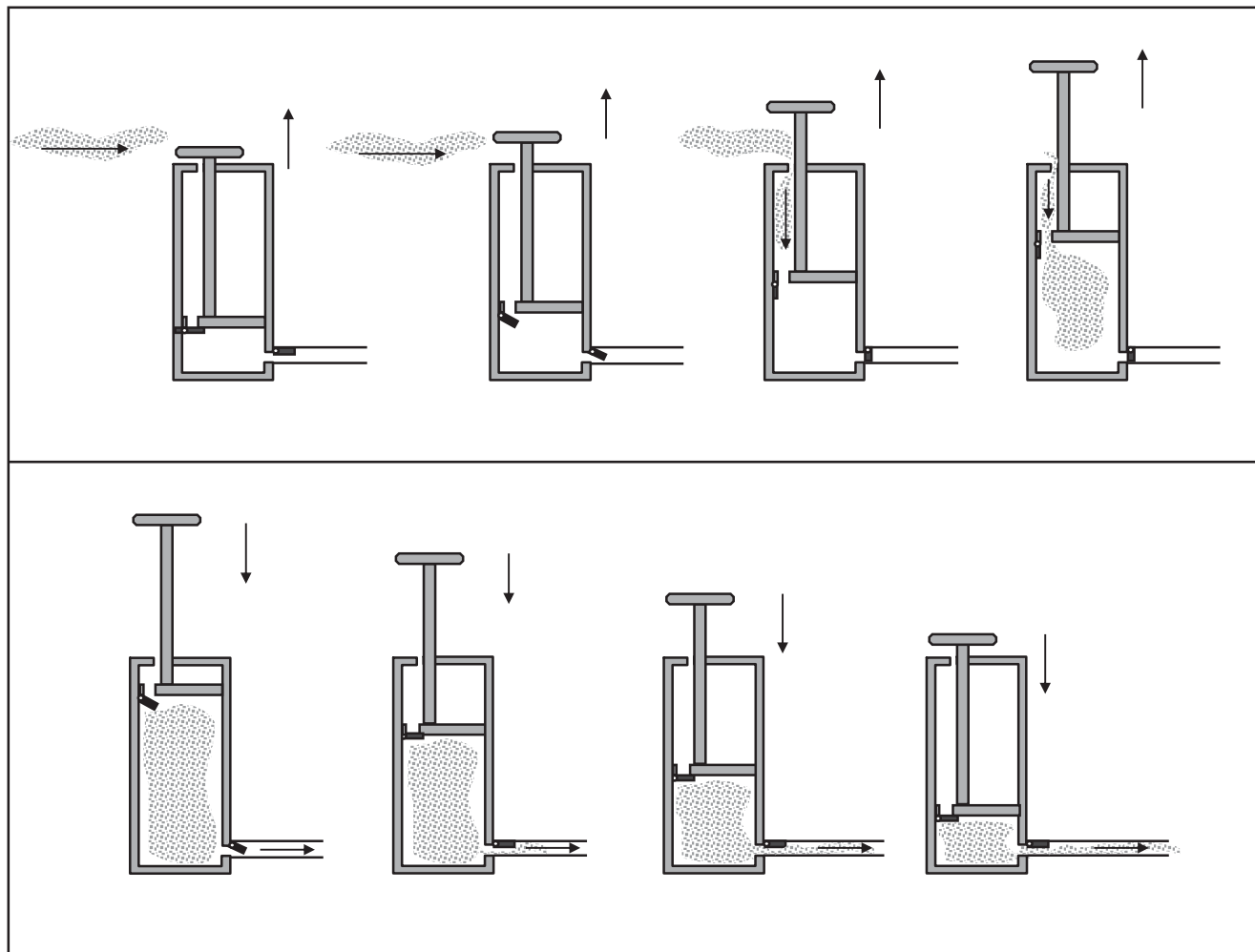

Figure 1.2 Frames from a pumps animation

"When the handle is pulled up, the piston moves up, the inlet valve opens, the outlet valve closes, and air enters the lower part of the cylinder."

"When the handle is pushed down, the piston moves down, the inlet valve closes, the outlet valve opens, and air moves out through the hose."
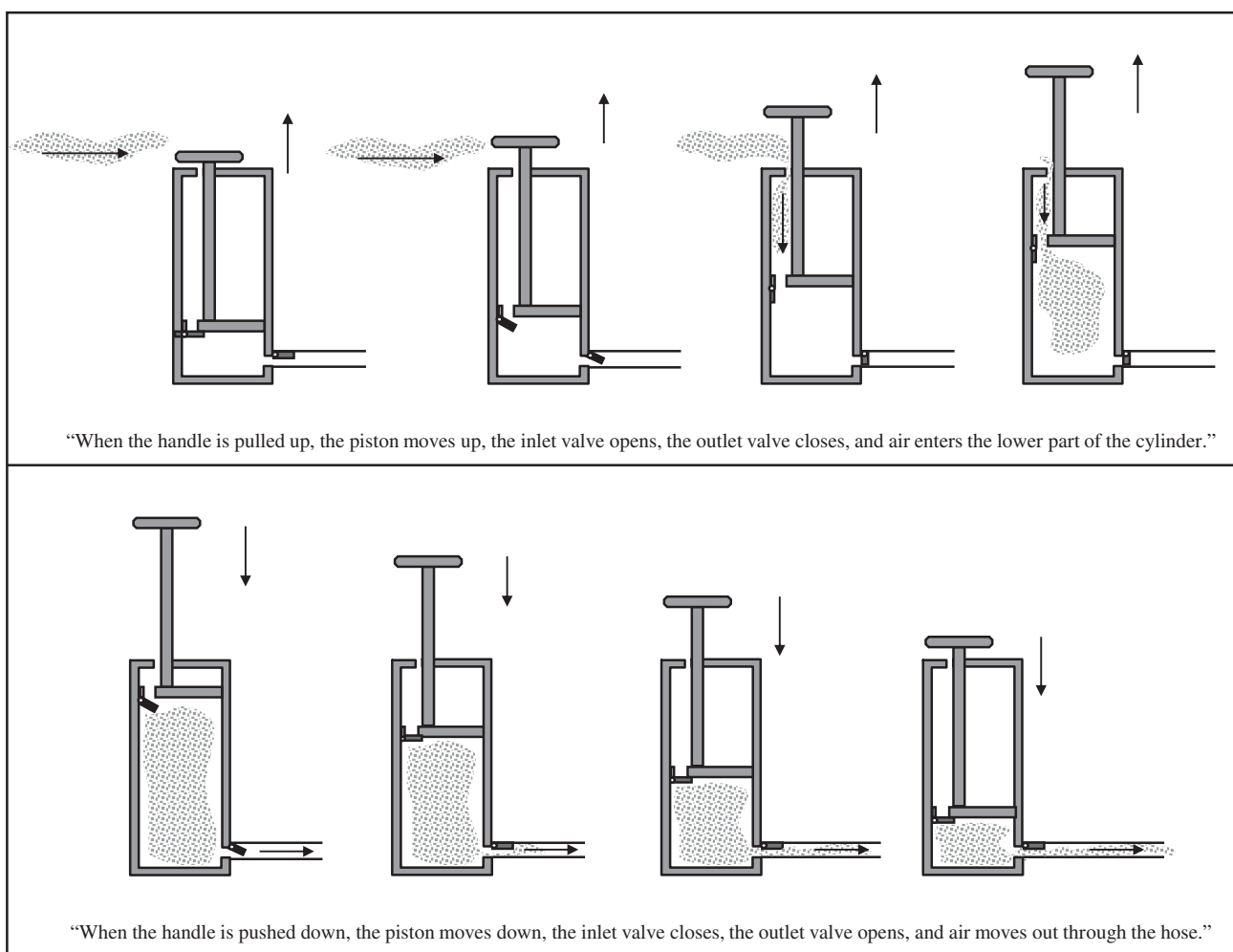
Figure 1.3 Frames from a pumps animation with corresponding narration

dominated education. Similarly, verbal learning has been the major focus of research on learning.

With the advent of powerful computer graphics and visualization technologies, instructors can supplement verbal modes of instruction with pictorial modes of instruction. Advances in computer technology have enabled an explosion in the availability of visual ways of presenting material, including large libraries of static images as well as applications for creating compelling dynamic images in the form of animations, video, and virtual reality. In light of the power of computer graphics, it may be useful to ask whether we should expand instructional messages beyond the purely verbal. The fundamental focus of this *Handbook* is on how to design instruction using words and pictures in ways that foster meaningful learning.

The case for multimedia learning is based on the idea that multimedia instructional messages should be designed in light of how the human mind works. Let's assume that humans have two information processing systems – one for verbal material and one for visual material, as described more fully in Part II of the *Handbook*. Let's also acknowledge that the dominant format for presenting instructional material has been mainly verbal. The rationale for multimedia presentations is that it takes advantage of the full capacity of humans for processing information. When we present material only in the verbal mode, we are ignoring the potential contribution of our capacity to also process material in the visual mode.

Why might two channels be better than one? Two possible explanations are the quantitative rationale and the qualitative rationale. The quantitative rationale is that more material can be presented on two channels than on one channel – just like more traffic can travel over two lanes than one lane. In the case of explaining how a bicycle tire pump works, for example, the steps in the process can be presented in words or can be depicted in illustrations. Presenting both is like presenting the material twice – giving the learner twice as much exposure to the explanation. While the quantitative rationale makes sense as far as it goes, we reject it mainly because it is incomplete. In particular, we take exception to the assumption that the verbal and visual channels are

equivalent; that is, that words and pictures are simply two equivalent ways for presenting the same material.

In contrast, the qualitative rationale is that words and pictures, while qualitatively different, can complement one another, and that human understanding is enhanced when learners are able to mentally integrate visual and verbal representations. As you can see, the qualitative rationale assumes that the two channels are not equivalent; words are useful for presenting certain kinds of material – perhaps representations that are more abstract and require more effort to translate – whereas pictures are more useful for presenting other kinds of material – perhaps more intuitive, more natural representations. In short, one picture is not necessarily the same as 1,000 words (or any number of words).

The most intriguing aspect of the qualitative rationale is that understanding occurs when learners are able to build meaningful connections between visual and verbal representations – such as being able to see how the words "the inlet valve opens" relate to the forward motion of the inlet valve in the cylinder of the pump. In the process of trying to build connections between words and pictures, learners are able to create a deeper understanding than from words or pictures alone. This idea is at the heart of the theories of multimedia learning described in Part II of the *Handbook*.

In summary, the rationale for the study of multimedia learning is that students may learn more deeply from words and pictures than from words alone. Thus, the motivation for this *Handbook* is to explore the proposal that adding pictures to words may promote understanding better than simply presenting words alone. However, not all pictures are equally effective. It is important to understand how best to incorporate pictures with words. Just because technologies are available that allow for state-of-the-art visualizations, this does not mean that instructors are well advised to use them. What we need is a research-based understanding of how people learn from words and pictures and how to design multimedia instruction that promotes learning.

TECHNOLOGY-CENTERED VERSUS LEARNER-CENTERED APPROACHES TO MULTIMEDIA LEARNING

Multimedia represents a potentially powerful learning technology – that is, a system for enhancing human learning. A practical goal of research on multimedia is to devise design principles for multimedia presentations. In addressing this goal, it is useful to distinguish between two approaches to multimedia design – a technology-centered approach and a learner-centered approach.

Technology-Centered Approaches

The most straightforward approach to multimedia design is technology-centered. Technology-centered approaches begin with the functional capabilities of multimedia and ask, "How can we use these capabilities in designing multimedia presentations?" The focus is generally on cutting-edge advances in multimedia technology, so technology-centered designers might focus on how to incorporate multimedia into emerging communications technologies such as wireless mobile access to the Internet or the construction of interactive multimedia representations in virtual reality. The kinds of research issues often involve media research, i.e., determining which technology is most effective for presenting information. For example, a media research issue is whether students learn as well from an online lecture – in which the student can see a lecturer in a window on the computer screen – as from a live lecture – in which the student is actually sitting in a classroom.

What's wrong with technology-centered approaches? A review of educational technologies of the twentieth century shows that the technology-centered approach generally fails to lead to lasting improvements in education (Cuban, 1986). For example, when the motion picture was invented in the early twentieth century hopes were high that this visual technology would improve education. In 1922, the famous inventor Thomas Edison predicted that "the motion picture is destined to revolutionize our educational system and that in a few years it will supplant largely, if not entirely, the use of textbooks" (cited in Cuban, 1986, p. 9). Like current claims for the power of visual media, Edison proclaimed that "it is possible to teach every branch of human knowledge with the motion picture" (cited in Cuban, 1986, p. 11). In spite of the grand predictions, a review of educational technology reveals that "most teachers used films infrequently in their classrooms" (Cuban, 1986, p. 17). From our vantage point well into the twenty-first century it is clear the predicted educational revolution in which movies would replace books has failed to materialize.

Consider another disappointing example that may remind you of current claims for the educational potential of the Internet. In 1932, Benjamin Darrow, founder of the Ohio School of the Air, proclaimed that radio could "bring the world to the classroom, to make universally available the services of the finest teachers, the inspiration of the greatest leaders …" (cited in Cuban, 1986, p. 19). His colleague, William Levenson, the director of the Ohio School of the Air, predicted in 1945 that a "radio receiver will be as common in the classroom as the blackboard" and "radio instruction will be integrated into school life" (cited in Cuban, 1986, p. 19). As we rush to wire our schools and homes for access to the educational content of the Internet, it is humbling to recognize what happened to a similarly-motivated movement for radio: "Radio has not been accepted as a full-fledged member of the educational community" (Cuban, 1986, p. 24).

Third, consider the sad history of educational television – a technology that combined the visual power of the motion picture with the worldwide coverage of radio.

By the 1950s, educational television was touted as a way to create a "continental classroom" that would provide access to "richer education at less cost" (Cuban, 1986, p. 33). Yet, a review shows that teachers used television infrequently if at all (Cuban, 1986).

Finally, consider the most widely acclaimed technological accomplishment of the twentieth century – computers. The technology that supports computers is different from film, radio, and television, but the grand promises to revolutionize education are the same. Like current claims for the mind-enhancing power of computer technology, during the 1960s computer tutoring machines were predicted to eventually replace teachers. The first large-scale implementation occurred under the banner of computer-assisted instruction (CAI) in which computers presented short frames, solicited a response from the learner, and provided feedback to the learner. In spite of a large financial investment to support CAI, sound evaluations showed that the two largest computer-based systems in the 1970s – PLATO and TICCIT – failed to produce better learning than traditional teacher-led instruction (Cognition and Technology Group at Vanderbilt, 1996).

What can we learn from the humbling history of the twentieth century's great educational technologies? Although different technologies underlie film, radio, television, and computer-assisted instruction, they all produced the same cycle. First, they began with grand promises about how the technology would revolutionize education. Second, there was an initial rush to implement the cutting-edge technology in schools. Third, from the perspective of a few decades later it became clear that the hopes and expectations were largely unmet.

What went wrong with these technologies that seemed poised to tap the potential of visual and worldwide learning? We attribute the disappointing results to the technology-centered approach taken by the promoters. Instead of adapting technology to fit the needs of human learners, humans were forced to adapt to the demands of cutting-edge technologies. The driving force behind the implementations was the power of the technology rather than an interest in promoting human cognition. The focus was on giving people access to the latest technology rather than helping people to learn through the aid of technology.

The cycle of technology-based approaches to education did not end with the turn of the century. For example, in the early twenty-first century we see many strong claims for the potential of digital games and simulations to replace traditional education, but there is a need to test these claims in rigorous scientific research (Mayer, 2014, 2019; see also Chapter 40).

Today, some commonly adopted cutting-edge technologies involve hand-held portable devices such as smartphones, tablets, e-readers, and controllers, as well as wearable devices such as smart watches, and devices for augmented and virtual reality. For example, school districts are told that the wave of the future requires purchasing one tablet for each student. Are we about to replicate the cycle of high expectations, large-scale implementation, and disappointing results in the realm of multimedia technology? In our opinion, the answer to that question depends on whether we continue to take a technology-centered approach. When we ask, "How can we give multimedia technology to students?" and when our goal is to "provide access to technology," we are taking a technology-centered approach with a 100-year history of failure.

Learner-Centered Approaches

Learner-centered approaches offer an important alternative to technology-centered approaches. Learner-centered approaches begin with an understanding of how the human mind works and ask, "How can we adapt multimedia to enhance human learning?" The focus is on using multimedia technology as an aid to human cognition. Research questions focus on the relation between design features and the human information processing system, such as comparing multimedia designs that place light or heavy loads on the learner's visual information processing channel. The premise underlying the learner-centered approach is that multimedia designs that are consistent with the way the human mind works are more effective in fostering learning than those that are not. This premise is the central theme of Part II of the *Handbook*, which lays out theories of multimedia learning.

Norman (1993, p. xi) eloquently makes the case for a learner-centered approach to technology design, which he refers to as human-centered technology: "Today we serve technology. We need to reverse the machine-centered point of view and turn it into a person-centered point of view: Technology should serve us." Norman's (1993, p. 12) vision of a learner-centered approach to technology design is that "technology ... should complement human abilities, aid those activities for which we are poorly suited, and enhance and help develop those for which we are ideally suited." In sum, as the twentieth century's most important new cognitive artifact, computer technology represents a landmark invention that has the potential to assist human cognition in ways that were previously not possible – but only if we take a learner-centered approach.

The differences between the technology-centered and learner-centered approaches to multimedia design are summarized in Table 1.2.

THREE METAPHORS OF MULTIMEDIA LEARNING: RESPONSE STRENGTHENING, INFORMATION ACQUISITION, AND KNOWLEDGE CONSTRUCTION

In making decisions about how to design or select a multimedia learning environment, you may be influenced by your underlying conception of learning. Table 1.3 compares three views of multimedia learning – *multimedia*

Table 1.2 Two approaches to the design of multimedia instruction

Design Approach	Starting Point	Goal	Issues
Technology-centered	Capabilities of multimedia technology	Provide access to information	How can we use cutting-edge technology in designing multimedia instruction?
Learner-centered	How the human mind works	Aid human cognition	How can we adapt multimedia technology to aid human cognition?

Table 1.3 Three metaphors of multimedia learning

Metaphor	Definition	Content	Learner	Teacher	Goal of multimedia
Response strengthening	Strengthening and weakening connections	Connections	Passive receiver of rewards and punishments	Dispenser of rewards and punishments	Exercise system
Information acquisition	Adding information to memory	Information	Passive receiver of information	Dispenser of information	Delivery system
Knowledge construction	Building a coherent mental structure	Knowledge	Active sense maker	Cognitive guide	Cognitive guidance system

learning as response strengthening, *multimedia learning as information acquisition*, and *multimedia learning as knowledge construction*. If you view multimedia learning as response strengthening, then multimedia is a feedback delivery system. If you view multimedia learning as information acquisition, then multimedia is an information delivery system. If you view multimedia learning as knowledge construction, then multimedia is a cognitive aid.

Multimedia Learning As Response Strengthening

According to the response strengthening view, learning involves increasing or decreasing the connection between a stimulus and a response. The underlying principle is that the connection is strengthened if a response is followed by reward and is weakened if the response is followed by punishment. This view entails assumptions about the nature of what is learned, the nature of the learner, the nature of the teacher, and the goals of multimedia presentations. First, learning is based on building connections, so "what is learned" is that a certain response is connected to a certain situation. Second, the learner's job is to make a response and receive feedback on the response; thus, the learner is a passive recipient of rewards and punishments. Third, the teacher's job – or, in some cases, the instructional designer's job – is to dispense rewards and punishments. Overall, the teacher controls the instructional episode by providing a prompt or question – such as, "What is the definition of multimedia learning?" – and then providing feedback on the answer given by the learner –

such as, "Yes, that's correct" or "No, you left out _____." Finally, the goal of multimedia instruction is to provide practice in exercising skills, that is, to act as a trainer. The underlying metaphor is that multimedia is an exercise system, that is, a system for practicing skills with feedback.

The response strengthening view reflects the first major theory of learning proposed by educational psychologists in the early 1900s – the law of effect (Thorndike, 1913). According to Thorndike's law of effect, if a response is followed by a satisfying state of affairs it will be more likely to occur under the same circumstances, and if a response is followed by a unsatisfying state of affairs it will be less likely to occur under the same circumstances. This straightforward principle has been a pillar of psychology and education for more than 100 years (Mayer, 2001), dominating the field through the 1950s. The law of effect was the guiding principle for many early instructional programs delivered by teaching machines in the 1960s. This view of learning can still be seen in multimedia environments that emphasize drill and practice, such as an online game that teaches arithmetic computation by giving the learner points for each correctly answered arithmetic problem.

What is wrong with the response strengthening view (or more accurately, the response strengthening and weakening view)? Our main objection is not that it is incorrect but rather that it is incomplete. Although certain cognitive skills (and motor skills, for that matter) can best be learned through drill and practice, the teaching of other kinds of knowledge – such as

concepts and strategies – may best be taught with other methods of instruction based on other views of learning. For example, when the goal of instruction is to foster meaningful learning reflected in the ability to solve transfer problems, drill and practice aimed at response strengthening may be too limited. Thus, the response strengthening view is appropriate for guiding the design of multimedia learning environments mainly when the goal of instruction is to help people learn specific skills. However, when the goal of instruction is to help people learn concepts and strategies that can be applied to new situations, the response strengthening view is not adequate.

Multimedia Learning As Information Acquisition

According to the information acquisition view, learning involves adding information to one's memory. As with the other views, the information acquisition view entails assumptions about the nature of what is learned, the nature of the learner, the nature of the teacher, and the goals of multimedia presentations. First, learning is based on information – an objective item that can be moved from place to place (such as from the computer screen to the human mind). Second, the learner's job is to receive information; thus, the learner is a passive being who takes in information from the outside and stores it in memory. Third, the teacher's job – or the multimedia designer's job – is to present information. Fourth, the goal of multimedia presentations is to deliver information as efficiently as possible. The underlying metaphor is that of multimedia as a delivery system; according to this metaphor, multimedia is a vehicle for efficiently delivering information to the learner.

The information acquisition view is sometimes called the *empty vessel* view because the learner's mind is seen as an empty container that needs to be filled by the teacher pouring in some information. Similarly, this is sometimes called the *transmission* view because the teacher transmits information to be received by the learner. Finally, this is sometimes called the *commodity* view because information is seen as a commodity than can be moved from one place to another.

What is wrong with the information acquisition view? If your goal is to help people learn isolated fragments of information, then we suppose nothing is wrong with the information acquisition view. However, when your goal is to promote understanding of the presented material, the information acquisition view is not very helpful. Even worse, it conflicts with the research base on how people learn complex material (Mayer, 2011, 2021). When people are trying to understand presented material – such as a lesson on how a bicycle tire pump works – they are not tape recorders who carefully store each word. Rather, humans focus on the meaning of presented material and interpret it in light of their prior knowledge.

Multimedia Learning As Knowledge Construction

In contrast to the information acquisition view, according to the knowledge construction view, multimedia learning is a sense-making activity in which the learner seeks to build a coherent mental representation from the presented material. Unlike information – which is an objective commodity that can be moved from one mind to another – knowledge is personally constructed by the learner and cannot be delivered in exact form from one mind to another. This is why two learners can be presented with the same multimedia message and come away with different learning outcomes. Second, according to the knowledge construction view, the learner's job is to make sense of the presented material; thus, the learner is an active sense-maker who interprets a multimedia presentation and tries to integrate the presented material into a coherent mental representation. Third, the teacher's job is to assist the learner in this sense-making process; thus, the teacher is a cognitive guide who provides needed guidance to support the learner's cognitive processing. Fourth, the goal of multimedia presentations is not only to present information, but also to provide guidance for how to process the presented information – that is for determining what to pay attention to, how to mentally organize it, and how to relate it to prior knowledge. Finally, the guiding metaphor is that of multimedia as a helpful communicator; according to this metaphor, multimedia is a sense-making guide, that is, an aid to knowledge construction.

Overall, we favor a knowledge construction view because it is more consistent with the research base on how people learn and because it is more consistent with our goal of promoting understanding of presented material. Rather than seeing the goal of multimedia presentations as exposing learners to vast quantities of information or exercising correct responses, our goal for multimedia is to help people develop an understanding of important aspects of the presented material. In short, the knowledge construction view offers a more useful conception of learning when the goal is to help people to understand and to be able to use what they learned.

THREE KINDS OF MULTIMEDIA LEARNING OUTCOMES: NO LEARNING, ROTE LEARNING, AND MEANINGFUL LEARNING

There are two major kinds of goals of learning – remembering and understanding. Remembering is the ability to reproduce or recognize the presented material, and is assessed by retention tests. The most common retention tests are recall – in which learners are asked to reproduce what was presented (such as writing down all they can remember for a lesson they read) – and recognition – in which learners are asked to select what was presented (as in a multiple choice question) or judge whether a given item was presented (as in a true–false question).

Thus, retention tests assess the quantity of learning – that is, how much was remembered.

Understanding is the ability to construct a coherent mental representation from the presented material; it is reflected in the ability to use the presented material in novel situations, and is assessed by transfer tests. In a transfer test, learners must solve problems that were not explicitly given in the presented material – that is, they must apply what they learned to a new situation. An example is an essay question that asks learners to generate solutions to a problem, which requires going beyond the presented material. Thus, transfer tests assess the quality of learning – that is, how well someone can use what they have learned. The distinction between remembering and understanding is summarized in Table 1.4. A major goal of the research presented in this *Handbook* is to promote understanding as well as retention.

Table 1.5 summarizes three kinds of learning outcomes: no learning, rote learning, and meaningful learning. The distinguishing feature of no learning is poor performance on retention and transfer. In this case, the learner lacks knowledge. The distinguishing pattern for rote learning outcomes is good retention and poor transfer. In this case, the learner has what can be called *fragmented knowledge* or *inert knowledge*, knowledge that can be remembered but cannot be used in new situations. In short, the learner has acquired a collection of *factoids* – isolated bits of information. Finally, meaningful learning is distinguished by good transfer performance as well as good retention performance. In this case, the learner's knowledge is organized into an integrated representation. Overall, the chapters in this *Handbook* examine design features of multimedia that foster meaningful learning, that is, ways of integrating words and pictures that foster transfer.

TWO KINDS OF ACTIVE LEARNING: BEHAVIORAL ACTIVITY VERSUS COGNITIVE ACTIVITY

What's the best way to promote meaningful learning outcomes? The answer rests in *active learning* – meaningful learning outcomes occur as a result of the learner's activity during learning. However, does active learning refer to what's going on with the learner's physical behavior – such as the degree of hands-on activity – or to what's going on in the learner's mind – such as the degree of integrative cognitive processing? In short, if the goal is to foster meaningful learning outcomes, should multimedia presentations be designed mainly to prime behavioral activity or cognitive activity?

Consider the following situation. Arlo is preparing for an upcoming test in meteorology. He sits in front of a computer and clicks on an interactive tutorial on lightning. The tutorial provides hands-on exercises in which he must fill in blanks by writing words. For example, on the screen appears the sentence: "Each year approximately _____ Americans are killed by lightning." He types in an answer, and the computer then provides the correct answer. In this case, Arlo is behaviorally active in that he is typing answers on the keyboard, but he may not be cognitively active in that he is not encouraged to make sense of the presented material.

In contrast, consider the case of Brianna, who is also preparing for the same upcoming meteorology test. Like Arlo, she sits in front of a computer and clicks on a tutorial about lightning; however, Brianna's tutorial is a short, narrated animation explaining the steps in the lightning formation. As she watches and listens, Brianna tries to focus on the essential steps in lightning formation and to organize them into a cause-and-effect chain. Wherever the multimedia presentation is unclear about why one step leads to another, Brianna uses her

Table 1.4 Two goals of multimedia instruction

Goal	Definition	Test	Example test item
Remembering	Ability to reproduce or recognize presented material	Retention	Write down all you can remember from the presentation you just studied
Understanding	Ability to use presented material in novel situations	Transfer	List some ways to improve the reliability of the device you just read about

Table 1.5 Three kinds of multimedia learning outcomes

Learning outcome	Cognitive description	Test performance	
		Retention	Transfer
No learning	No knowledge	Poor	Poor
Rote learning	Fragmented knowledge	Good	Poor
Meaningful learning	Integrated knowledge	Good	Good

prior knowledge to help create an explanation for herself – which Chi, Bassok, Lewis, Reimann, and Glaser (1989) call a *self-explanation* (see also Chapter 32). For example, when the narration says that positively charged particles come to the surface of the earth, Brianna mentally creates the explanation that opposite charges attract. In this scenario, Brianna is behaviorally inactive because she simply sits in front of the computer; however, she is cognitively active because she is actively trying to make sense of the presentation.

Which type of active learning promotes meaningful learning? Research on learning shows that meaningful learning depends on the learner's cognitive activity during learning rather than on the learner's behavioral activity during learning. You might suppose that the best way to promote meaningful learning is through hands-on activity, such as a highly interactive multimedia program. However, behavioral activity per se does not guarantee cognitively active learning; it is possible to engage in hands-on activities that do not promote active cognitive processing – such as in the case of Arlo or many highly interactive computer games. You might suppose that presenting material to a learner is not a good way to promote active learning because the learner appears to sit passively. In some situations, your intuition would be right – presenting a long, incoherent, and boring lecture or textbook chapter is unlikely to foster meaningful learning. However, in other situations, such as the case of Brianna, learners can achieve meaningful learning in a behaviorally inactive environment such as a multimedia instructional message. Our point is that well-designed multimedia instructional messages can promote active cognitive processing in learners, even when learners seem to be behaviorally inactive.

WHAT IS THE RESEARCH BASE FOR MULTIMEDIA LEARNING?

Although research on verbal learning has a long and fruitful history in psychology and education, corresponding research on multimedia learning is now beginning to flourish. This third edition of *The Cambridge Handbook of Multimedia Learning* remains the world's first and most comprehensive summary of research on multimedia learning. In order to organize the large research base in multimedia learning, the *Handbook* is divided into eight parts.

Part I: Background sets the stage for understanding the field of multimedia learning with chapters that provide an introduction to key concepts (Chapter 1, by Mayer and Fiorella), provide historical context (Chapter 2, by Camp, Surma, and Kirschner), expose mistaken principles of multimedia learning, that is, principles that are commonly accepted but for which supporting evidence is lacking (Chapter 3, by Clark, Feldon, and Jeong), and give an overview of research methods used to study multimedia learning (Chapter 4, by Jarodzka). Overall, Part I provides

background needed for understanding the contributions of research on multimedia learning explored in subsequent sections.

Part II: Theoretical Foundations helps ground the field of multimedia learning within evidence-based theories of multimedia learning with chapters that describe relevant theories that have had the greatest impact on research: Mayer's cognitive theory of multimedia learning (Chapter 5, by Mayer), Sweller's cognitive load theory (Chapter 6, by Paas and Sweller), Schnotz's integrative model of text and picture comprehension (Chapter 7, by Schnotz), and van Merriënboer's four-component instructional design model for multimedia learning (Chapter 8, by Kester and van Merriënboer). In addition to these foundational theories that focus on the cognitive processes in multimedia learning, this section contains chapters on two kinds of processes that need to be integrated into any complete account of multimedia learning: motivational processes (Chapter 9, by Schrader, Kalyuga, and Plass) and metacognitive processes (Chapter 10, by Azevedo and Dever). Overall, Part II gives a detailed look at the mechanisms by which people learn from multimedia instructional messages, which contributes to and is shaped by the research findings described in subsequent sections.

Part III: Basic Principles of Multimedia Learning focuses on the foundational principles of multimedia learning:

multimedia principle: learning from words and pictures is more effective than learning from words alone (Chapter 11, by Mayer), related to research on the processes involved in learning from multiple representations (Chapter 12, by Ainsworth).

expertise reversal principle: instructional methods of multimedia instruction that are effective for less experienced learners may not be effective for more experienced learners and vice versa (Chapter 13, by Kalyuga).

Overall, Part III introduces two foundational ideas: (1) multimedia instruction is more effective than purely verbal instruction and (2) there are individual differences in how people benefit from instructional design features intended to improve multimedia instruction.

Part IV: Principles for Reducing Extraneous Processing in Multimedia Learning contains chapters that explore the research evidence concerning basic principles for how to minimize extraneous processing in multimedia learning – that is, cognitive processing that does not support the instructional goal. These principles for reducing extraneous processing are:

coherence principle: people learn better from multimedia instructional messages when extraneous words and images are excluded (Chapter 14, by Fiorella and Mayer).

signaling (or cueing) principle: people learn better when cues are added that highlight or spotlight the key

information in a multimedia lesson and its organization (Chapter 17, by van Gog), similar to Fiorella and Mayer's *signaling principle* (in Chapter 14).

redundancy principle: people learn better when the same information is not presented in more than one format (Chapter 16, by Kalyuga and Sweller), similar to Fiorella and Mayer's *redundancy principle* (in Chapter 14).

split-attention principle: people learn better when words and pictures are physically and temporally integrated (in Chapter 15, by Ayres and Sweller), similar to the *spatial contiguity principle* and *temporal contiguity principle* (in Chapter 14, by Fiorella and Mayer).

worked example principle: people learn better from seeing worked-out examples in initial learning of cognitive skills (Chapter 18, by Renkl).

Overall, there is a solid research base explicating instructional design techniques for reducing extraneous processing during multimedia learning, which was the original focus of much multimedia learning research (Mayer, 2021).

Part V: Principles for Managing Essential Processing in Multimedia Learning contains chapters that explore the research base for instructional design principles aimed at managing essential cognitive processing in multimedia learning – that is, processing aimed at mentally representing the presented material. These principles for managing essential processing include:

segmenting principle: people learn better when a multimedia message is presented in learner-paced segments rather than as a continuous unit (Chapter 19, by Mayer and Fiorella).

pre-training principle: people learn better from a multimedia message when they know the names and characteristics of the main concepts (Chapter 19, by Mayer and Fiorella).

modality principle: people learn better from a multimedia message when the words are spoken rather than written (Chapter 19, by Mayer and Fiorella; Chapter 20, by Castro-Alonso and Sweller), related to the *transient information principle* (Chapter 21, by Jiang and Sweller).

Overall, there is a solid research base explicating instructional design techniques for managing essential processing during multimedia learning, especially concerning the role of using spoken versus printed text.

Part VI: Principles Based on Social and Affective Features of Multimedia Learning examines evidence-based principles aimed at fostering generative cognitive processing during multimedia learning (i.e., processing based on exerting effort to make sense of the material). These principles for promoting generative processing include:

personalization principle: people learn better when the words of a multimedia presentation are in conversational style rather than formal style (Chapter 22, by Fiorella and Mayer).

voice principle: people learn better when the words in a multimedia lesson are spoken in an appealing human voice rather than a machine voice (Chapter 22, by Fiorella and Mayer).

embodiment principle: people learn better when onscreen agents display humanlike gestures and movements (Chapter 22, by Fiorella and Mayer; Chapter 23, by Fiorella).

image principle: people do not necessarily learn better when the speaker's static image is on the screen (Chapter 22, by Fiorella and Mayer).

immersion principle: people do not necessarily learn better with higher immersion media (such as immersive virtual reality) than lower immersion media (such as onscreen video) (Chapter 24, by Makransky).

collaboration principle: people learn better in groups than individually (Chapter 25, by Janssen, Kirschner, and Kirschner).

animation principle: people learn better from dynamic graphics than static graphics (Chapter 26, by Lowe, Schnotz, and Boucheix).

emotional design principle: people learn from multimedia lessons involving elements with facial expressions and warm colors (Chapter 27, by Plass and Hovey).

Overall, this section of the *Handbook* explores emerging research on the effectiveness of design elements intended to create social partnership or prime an affective reaction in learners.

Part VII: Principles Based on Generative Activity in Multimedia Learning focuses on the research evidence for design principles based on prompting the learner to engage in *generative learning activities*, that is, behaviors that the learner engages in during learning with the intention of improving learning. As in Part VI, these principles are intended to foster generative processing during learning. They include:

generative activity principle: people learn better when they are prompted to engage in generative learning strategies during learning (such as summarizing, mapping, drawing, imagining, self-explaining, self-testing, explaining to others, and enacting) (Chapter 28, by Fiorella and Mayer).

mapping principle: people learn better when they are encouraged to create concept maps, knowledge maps, or graphic organizers during learning (Chapter 29, by Adesope, Nesbit, and Sundararajan; Chapter 28, by Fiorella and Mayer).

drawing principle: people learn better when they are prompted to create drawings as they read explanative text (Chapter 30, by Leutner and Schmeck; Chapter 28, by Fiorella and Mayer).

imagination principle: people learn better when they are prompted to imagine drawings as they read explanative

text (Chapter 31, by Leopold; Chapter 28, by Fiorella and Mayer).

self-explanation principle: people learn better when they are encouraged to generate self-explanations during learning (Chapter 32, by Chi; Chapter 28, by Fiorella and Mayer).

guided-discovery principle: people learn better when guidance is incorporated into discovery-based multimedia environments (Chapter 33, by de Jong).

feedback principle: people learn better from multimedia lessons when they receive explanative feedback on their performance (Chapter 34, by Johnson and Marraffino).

learner control principle: people do not necessarily learn better when they have more control of the selection and organization of the material (Chapter 35, by Scheiter).

self-management principle: people learn better when they are prompted to exercise control over their learning processes (Chapter 36, by Zhang, de Koning, Agostinho, Tindall-Ford, Chandler, and Paas).

Overall, this section shows the current state of emerging research on the effectiveness of prompts to engage in generative learning activities.

Part VIII: Multimedia Learning with Media takes a cross-cutting approach by looking at what the research has to say about increasing the effectiveness of multimedia learning with media such as online cognitive tutors (Chapter 37, by Koedinger and Aleven), animated pedagogical agents (Chapter 38, by Wang, Li, and Zhao), simulations (Chapter 39, by Lajoie), computer games (Chapter 40, by Mayer), instructional video (Chapter 41, by Fiorella), virtual reality (Chapter 42, by Parong), visual displays (Chapter 43, by McCrudden and Van Meter), multiple documents (Chapter 44, by Rouet and Britt), e-courses (Chapter 45, by Clark), and multimedia assessment (Chapter 46, by Lindner). Overall, this section summarizes some of the emerging research on multimedia learning afforded by various media platforms. Overall, each *Handbook* chapter in Parts III–VIII is intended to showcase the research base in a sub-area of multimedia learning, note its limitations, pinpoint practical and theoretical implications, and offer suggestions for future research.

WHAT'S TRENDING IN THE THIRD EDITION?

A unifying framework for this third edition of *The Cambridge Handbook of Multimedia Learning* is to separate the research base on principles for multimedia instruction into categories based on three instructional goals: reducing extraneous processing (examined in Part IV), managing essential processing (examined in Part V), and fostering generative processing (examined in Parts VI and VII). Although the overarching goal remains the same as in previous editions (i.e., to take an evidence-based approach to the design of multimedia instruction), we

see six major trends in this third edition of the *Handbook* concerning:

what works: this *Handbook* highlights substantial increases in the quantity and quality of the empirical research on design principles in all three areas (examined in Parts IV–VII) with an especially impressive increase in research aimed at design principles for fostering generative processing during multimedia learning, including research on voice, embodiment, immersion, emotional design, and generative learning activities (examined in Parts VI and VII).

when it works: this *Handbook* also highlights increases in our understanding of moderator variables that pinpoint the boundary conditions of multimedia design principles including for whom they work, and for which kinds of materials, subject domains, learning objectives, and learning situations they work.

where it works: this *Handbook* reflects a broadening of the learning venue by examining the design of multimedia instruction with various media, ranging from instructional video to immersive virtual reality to pedagogical agents (examined in Part VIII).

how it works: this *Handbook* also reflects refinements in theories of multimedia learning, particularly attempts to broaden cognitive theories to incorporate motivational, affective, social, and metacognitive processes.

can it work in practical contexts: we see the continuation of emerging efforts to determine the degree to which multimedia design principles apply beyond the lab to school-based and adult training scenarios, including delayed testing and longer-term instructional periods. As suggested in the chapter on e-courses (Chapter 45) the worldwide pandemic has drawn attention to the practical need for effective remote education.

how do we know it works: Although most studies still focus on self-report surveys to assess cognitive processing during learning and immediate posttests to assess learning outcomes, we see an increasing number of studies that use additional assessment methods for learning processes including metrics derived from log files and eye-tracking measures, and we are beginning to see a few studies using biometric and brain-based measures. In addition, although most studies still rely on t-tests or analysis of variance to examine group differences, we see an increasing number of studies that use more advanced statistical methods such as mediational analysis and structural equation modeling (SEM) to investigate the role of moderating variables that may suggest boundary conditions for effects or mediating variables that suggest processes that occur during learning.

SUMMARY

In summary, this *Handbook* explores how to promote multimedia learning – that is, learning from words and pictures. In 46 chapters, the book takes an evidence-based

approach by examining what research has to say about how to design multimedia learning environments that help people learn. The approach taken in this *Handbook* is learner-centered rather than technology-centered and seeks to foster meaningful learning rather than rote learning. Overall, the *Handbook* examines the evidence for 30 principles of multimedia instructional design and explores their application in a variety of contexts, ranging from computer-based presentations to educational games to tutoring systems. The principles are organized into three categories based on the instruction goal: principles for reducing extraneous processing, principles for managing essential processing, and principles for fostering generative processing. Compared to the first edition and second edition of the *Handbook*, this third edition reflects a substantial growth in the research base underlying each principle, a better understanding of the boundary conditions for each principle, a broadening of the media under study, a broadening of theories of multimedia learning, increased interest in multimedia learning beyond the lab, and an increase in the arsenal of research tools.

REFERENCES

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145–182.

Cognition and Technology Group at Vanderbilt. (1996). Looking at technology in context: A framework for understanding technology in education. In D. Berliner, & R. C. Calfee (eds.), *Handbook of Educational Psychology* (pp. 807–840). New York: Macmillan.

Comenius, J. A. (1887). *Orbis Pictus*. Syracuse, NY: Bardeen. [Reproduced version.]

Cuban, L. (1986). *Teachers and Machines: The Classroom Use of Technology Since 1920*. New York: Teachers College Press.

Mayer, R. E. (2001). Changing conceptions of learning: A century of progress in the scientific study of education. In L. Corno (ed.), *Education across a Century: The Centennial Volume. One Hundredth Yearbook of the National Society for the Study of Education* (pp. 34–75). Chicago, IL: University of Chicago Press.

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

Mayer, R. E. (2014). *Computer Games for Learning: An Evidence-Based Approach*. Cambridge, MA: MIT Press.

Mayer, R. E. (2019). Computer games in education. *Annual Review of Psychology*, *70*, 531–549.

Mayer, R. E. (2021). *Multimedia Learning* (3rd ed.). New York: Cambridge University Press.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, *83*, 484–490.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, *84*, 444–452.

Norman, D. A. (1993). *Things That Make Us Smart*. Reading, MA: Addison-Wesley.

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. Oxford: Oxford University Press.

Paivio, A. (2007). *Mind and Its Evolution: A Dual Coding Theoretical Approach*. Mahwah, NJ: Erlbaum.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction*, *13*, 141–156.

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

Thorndike, E. L. (1913). *Educational Psychology*. New York: Columbia University Press.

van Merrienboer, J. J. G., & Kirschner, P. A. (2013). *Ten Steps to Complex Learning: A Systematic Approach to Four-Component Instructional Design* (2nd ed.). New York: Routledge.