Educational Psychology Review, Vol. 14, No. 1, March 2002 (© 2002)

Animation as an Aid to Multimedia Learning

Richard E. Mayer[1,3] and Roxana Moreno[2]

How can animation be used to promote learner understanding of scientific and mathematical explanations? In this review, we examine the role of animation in multimedia learning (including multimedia instructional messages and microworld games), present a cognitive theory of multimedia learning, and summarize our program of research, which has yielded seven principles for the use of animation in multimedia instruction. These include the multimedia principle (present animation and narration rather than narration alone), spatial contiguity principle (present on-screen text near rather than far from corresponding animation), temporal contiguity principle (present corresponding animation and narration simultaneously rather than successively), coherence principle (exclude extraneous words, sounds, and video), modality principle (present animation and narration rather than animation and on-screen text), redundancy principle (present animation and narration rather than animation, narration, and on-screen text), and personalization principle (present words in conversational rather than formal style). Animation can promote learner understanding when used in ways that are consistent with the cognitive theory of multimedia learning.

KEY WORDS: multimedia; animation; technology; science education; problem solving.

Multimedia instructional environments are widely recognized to hold great potential for improving the way that people learn (Mayer, 1999, in press; Sweller, 1999; van Merrienboer, 1997). In multimedia instructional environments, learners are exposed to material in verbal (such as on-screen text or narration) as well as pictoral form (including static materials such as

[1]University of California, Santa Barbara, California.
[2]University of New Mexico, New Mexico.
[3]Correspondence should be addressed to Richard E. Mayer, Department of Psychology, University of California, Santa Barbara, California 93106; e-mail: mayer@psych.ucsb.edu.

photos or illustrations, and dynamic materials such as video or animation).[4] Although verbal forms of presentation have long dominated education, there is encouraging evidence that student understanding can be enhanced by the addition of visual forms of presentation (Mayer, 1999, in press; Sweller, 1999).

WHAT IS THE ROLE OF ANIMATION IN MULTIMEDIA LEARNING?

One of the most exciting forms of pictoral presentation is animation. Animation refers to a *simulated motion picture* depicting movement of drawn (or simulated) objects. The main features of this definition are as follows: (1) picture – an animation is a kind of pictorial representation; (2) motion – an animation depicts apparent movement; and (3) simulated – an animation consists of objects that are artificially created through drawing or some other simulation method. In contrast, video refers to a motion picture depicting movement of real objects. Similarly, an illustration is a static picture of drawn (or simulated) objects whereas a photo is a static picture of real objects.

When used mainly as a form of entertainment, an animation can be called a cartoon, but in this review we focus on the potential of animation as an educational tool. Does animation promote learning? Do students learn more from animation than from other modes of presentation? Should we increase the use of animation in educational programs? These questions fit within a classical tradition of media research, in which the goal is to determine whether students learn better with one medium compared with another. However, media researchers have concluded that media research questions such as these are largely fruitless (Clark, 1994; Kozma, 1994; Ross, 1994; Salomon, 1979/1994).

The consensus among media researchers is that animation may or may not promote learning, depending on how it used. For these reasons the search for media effects has been called off. In its place is a search for the conditions under which various media, such as animation, affect the learning process. Taking a learner-centered approach, we aim to understand how animation can be used in ways that are consistent with how people learn. Instead of asking, "does animation improve learning?" we ask "when and how does animation affect learning?"

[4]Multimedia can be defined in terms of sensory modalities (e.g., visual vs. auditory), representational modes (e.g., pictorial vs. verbal), or delivery media (e.g., screens vs. speakers). In this paper, we define multimedia in terms of modes and modalities (e.g., visual/pictorial vs. auditory/verbal), so animation is processed in the visual/pictorial channel and narration is processed in the auditory/verbal channel.

Consider the following scenario: Alice is writing a report on film making. To gather research information, she goes to her on-line encyclopedia and clicks on “animation.” On the computer screen there appears a window containing a 250-word explanation of how animated cartoons are created and another window containing a 30-s animation depicting the process of creating an animated cartoon. She reads the words, having to scroll down several times, and then clicks on the “start” button to view the animation. This is an example of multimedia learning because the material is presented in pictoral and verbal forms, namely animation and on-screen text. In this case animation is used in the context of a *multimedia instructional message*—a multimedia presentation intended to explain something to a learner.

In our research, we have created four multimedia instructional messages consisting of narrated animations that explain how lightning storms form, how pumps work, how car brakes work, and how human lungs work (see Mayer, 1997, 1999, in press). The messages are short and focused, ranging from 30 to 180 s. For example, the multimedia instructional message for lightning consists of an animation that depicts the steps in the process of lightning formation (e.g., cool moist air moving over a warmer surface, moist air rising to form a cloud, and so on), and a corresponding narration that describes the steps in spoken words. To test learner understanding, we ask students to write answers to a series of four transfer questions such as, “What could be done to decrease the intensity of a lightning storm?” or “Suppose you see clouds in the sky but no lightning. Why not?” We score the transfer test by tallying the number of acceptable answers the student generated across all four problems, based on a list of acceptable answers for each problem.

Consider a second scenario: Beth is interested in playing a new physics game, so she calls up an educational game called “Click World” and clicks on “level 1.” She guides a ball through a maze by pressing buttons and moving a joystick and receives verbal feedback from a “computer coach” who talks to her as she plays. This is an example of multimedia learning because the material is presented in visual and verbal forms, namely animation and speech. In this case animation is used in the context of a *microworld game*, a simulated version of a real situation. In microworld games, students may interact with an animation-based simulation such as a game intended to teach scientific or mathematical concepts (Dunbar, 1993; Moreno and Mayer, 2000; Moreno *et al.*, 2000; Rieber, 1990; White, 1993).

In our research, we have studied how students learn in two microworld games: a mathematics game in which elementary-school children learn how to add and subtract signed numbers by moving a bunny along a number line (Moreno and Mayer, 2000), and a botany game in which high-school and college students design plants to survive in various environments (Moreno *et al.*, 2000). For example, in the botany game students interact with an animated pedagogical agent named “Herman the Bug” who takes them on a space ship to a new planet. The planet has certain environmental features such as low sunlight or heavy rain, and the student is asked to design a plant that would survive there—including selecting the appropriate type of roots, stem, and leaves (with eight alternatives for each). Herman provides help by showing an animation of plant growth while verbally describing how plants grow. Students travel to several different planets, get to design a plant for each, and then get to see if it survives. As a test of understanding, our transfer test involves designing plants for new environments and telling which kinds of environments a given plant is best suited for. The score is based on the number of correct answers the student gives across the transfer problems.

In this review, we focus on the role of animation in multimedia instructional messages and in microworld games, and we ask, “How should animation be presented to promote understanding of a multimedia explanation?” We begin by examining two views of multimedia learning, and then explore research on several design principles involving the role of animation in instruction.

HOW DO PEOPLE LEARN FROM WORDS AND PICTURES?

In designing multimedia presentations involving animation, instructional designers base their decisions on a theory of how students learn. In this section, we explore two competing views of how students learn from words and pictures.

Information Delivery Theory of Multimedia Learning

A straightforward theory is that learning involves adding information to one’s memory (see Mayer, 1996, in press). According to this theory, the computer is a system for delivering information to learners. The instructional designer’s role is to present information (e.g., as words or pictures, or both) and the learner’s role is to receive the information. For example, when an explanation is presented in words (such as narration) the learner can store the information in memory. Adding pictures (such as animation) should have no effect on what is learned if the pictures contain the same information as the words. Thus, according to this strict version of the information delivery theory, multimedia presentations should not result in better learning than single-medium presentations. However, if some learners prefer visual presentations and others prefer verbal presentations, then a multimedia presentation would be effective in delivering information effectively to both kinds of learners. In this way, learners could select the delivery route they prefer. Thus, according to a lenient version of the information delivery theory, multimedia presentations should result in better learning than single medium presentations.

Cognitive Theory of Multimedia Learning

An alternative idea is that meaningful learning occurs when students mentally construct coherent knowledge representations (see Mayer, 1996, in press). The cognitive theory of multimedia learning is based on three assumptions suggested by cognitive research: (1) dual-channel assumption – the idea that humans have separate channels for processing visual/pictorial representations and auditory/verbal representations (Baddeley, 1998; Paivio, 1986); (2) limited capacity assumption – the idea that only a few pieces of information can be actively processed at any one time in each channel (Baddeley, 1998; Sweller, 1999); and (3) active processing – the idea that meaningful learning occurs when the learner engages in cognitive processes such as selecting relevant material, organizing it into a coherent representation, and integrating it with existing knowledge (Mayer, in press; Wittrock, 1974).

Figure 1 summarizes the cognitive theory of multimedia learning. Narration enters via the ears, so the learner selects some of the words for further processing in the verbal channel, organizes the words into a cause-and-effect chain, and integrates it with the visual material and prior knowledge. Animation enters via the eyes, so the learner selects some of the images for further processing in the visual channel, organizes the images into a cause-and-effect chain, and integrates it with the verbal material and prior knowledge. According to this theory, the cognitive process of integrating is most likely to occur when the learner has corresponding pictoral and verbal representations in working memory at the same time. Instructional conditions that promote these processes are most likely to result in meaningful learning. As you can see, this theory predicts that multimedia presentations (such as narrated animation) are more likely to lead to meaningful learning than single-medium presentations.

HOW SHOULD ANIMATION BE USED WITHIN MULTIMEDIA PRESENTATIONS?

For the past decade, we and our colleagues at the University of California, Santa Barbara (UCSB), have been examining the conditions under which animation promotes learner understanding. In this section, we

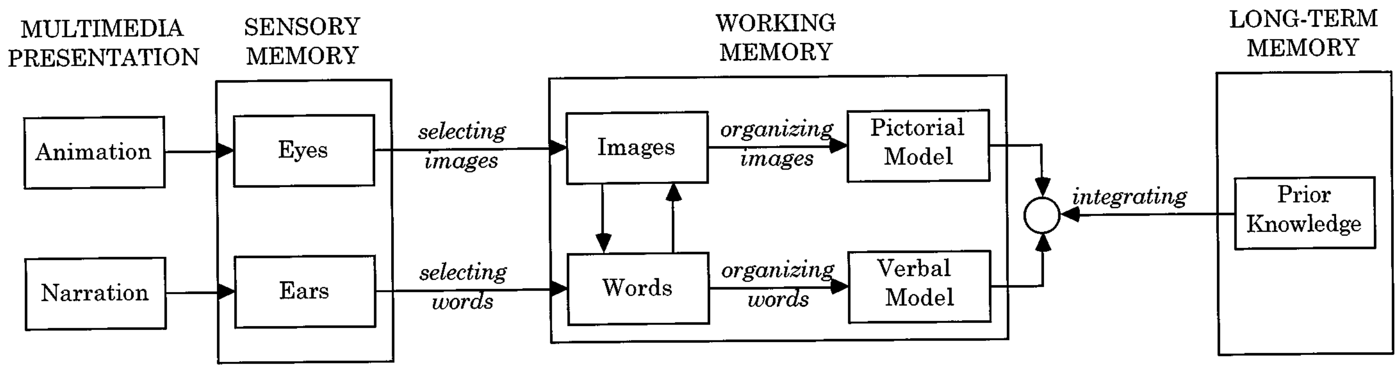

Fig. 1. A cognitive theory of multimedia learning.

summarize the fruits of our research, namely a collection of seven research-based principles for the design of multimedia presentations involving animation. Table I lists each of the principles, along with a summary of supporting research studies. For each study, we compared the problem-solving transfer performance of college students who learned with and without a certain condition (such as temporal coordination of animation and narration). Specifically, we computed an effect size as the difference between the mean transfer scores of the groups divided by the standard deviation of the group that did not receive the condition. We used transfer as our measure of meaningful learning (or learner understanding) because transfer taps students' ability to use what they have learned in new situations.

Multimedia Principle

The first principle is that students learn more deeply from animation and narration than from narration alone. The theoretical rationale for this principle is that students are better able to build mental connections between corresponding words and pictures when both are presented (i.e., animation and narration) than when only one is presented (i.e., narration) and the learner must mentally create the other. In each of four experiments, adding a pictoral explanation (i.e., animation) to a verbal one (i.e., narration) resulted in a substantial improvement in learners' problem-solving transfer performance. The median effect size was 1.74, indicating a strong and consistent effect. In these studies, animation enhanced student understanding of scientific explanations of how pumps work, how brakes work, or how to add and subtract signed numbers. In short, we have consistent evidence for the multimedia principle that words and pictures are better for promoting learner understanding than are words alone. Not all animations are equally effective in promoting understanding in learners, so each of the next six principles focuses on the difference between effective and ineffective uses of animation.

Spatial Contiguity Principle

The second principle is that students learn more deeply when on-screen text is presented next to the portion of the animation that it describes than when on-screen text is presented far from the corresponding action in the animation. The theoretical rationale is that learners are better able to build mental connections between corresponding words and pictures when they are near each other on the screen; in contrast, when they are not near each

Table I. Seven Principles of Multimedia Learning

Principle and comparison source	Materials	Effect size	Median effect size	Number of positive tests
1. Multimedia principle: Deeper learning from animation and narration than from narration alone			1.73	4 of 4
Mayer and Anderson (1992, Exp. 1)	Pumps	1.90		
Mayer and Anderson (1992, Exp. 2)	Brakes	1.67		
Mayer and Anderson (1991, Exp. 2a)	Pumps	2.43		
Moreno and Mayer (1999b, Exp. 1)	Math game	0.47		
2. Spatial contiguity principle: Deeper learning when corresponding text and animation are presented near rather than far from each other on the screen			0.48	1 of 1
Moreno and Mayer (1999a, Exp. 1)	Lightning	0.48		
3. Temporal contiguity principle: Deeper learning when corresponding narration and animation are presented simultaneously rather than successively			1.30	8 of 8
Mayer *et al*. (1999, Exp. 1)	Lightning	1.96		
Mayer *et al*. (1999, Exp. 2)	Brakes	1.27		
Mayer and Sims (1994, Exp. 1)	Brakes	0.83		
Mayer and Sims (1994, Exp. 2)	Lungs	1.60		
Mayer and Anderson (1992, Exp. 1)	Pumps	1.61		
Mayer and Anderson (1992, Exp. 2)	Brakes	1.33		
Mayer and Anderson (1991, Exp. 1)	Pumps	1.00		
Mayer and Anderson (1991, Exp. 2a)	Pumps	1.05		
4. Coherence principle: Deeper learning when extraneous narration, sounds, and video are excluded rather than included			0.90	5 of 5
Mayer *et al.* (2001, Exp. 1)	Lightning	0.55		
Moreno and Mayer (2000a, Exp. 1)	Lightning	1.56		
Moreno and Mayer (2000a, Exp. 2)	Brakes	0.90		
Mayer *et al.* (2001, Exp. 3)	Lightning	0.86		
Mayer *et al.* (2001, Exp. 4)	Lightning	1.03		
5. Modality principle: Deeper learning from animation and narration than from animation and on-screen text			1.17	6 of 6
Moreno and Mayer (1999a, Exp. 1)	Lightning	1.06		
Moreno and Mayer (1999a, Exp. 2)	Lightning	1.28		
Mayer and Moreno (1998, Exp. 1)	Lightning	1.68		
Mayer and Moreno (1998, Exp. 2)	Brakes	0.94		
Moreno *et al.* (2000, Exp. 1)	Botany game	0.89		
Moreno *et al.* (2000. Exp. 2)	Botany game	1.37		
6. Redundancy principle: Deeper learning from animation and narration than from animation, narration, and on-screen text			0.77	2 of 2
Mayer *et al.* (2001, Exp. 1)	Lightning	0.66		
Mayer *et al.* (2001, Exp. 2)	Lightning	0.88		
7. Personalization principle: Deeper learning when narration or on-screen text is conversational rather than formal			1.55	5 of 5
Moreno and Mayer (2000b, Exp. 1)	Lightning	1.00		
Moreno and Mayer (2000b, Exp. 2)	Lightning	1.60		
Moreno and Mayer (2000b, Exp. 3)	Botany game	1.55		
Moreno and Mayer (2000b, Exp. 4)	Botany game	1.58		
Moreno and Mayer (2000b, Exp. 5)	Botany game	0.89		

other, learners must waste limited cognitive capacity in searching for the portion of the animation that corresponds to the presented text. In the one experiment that we conducted, students who received on-screen text presented next to the corresponding event in the animation performed better on problem-solving transfer than students who received on-screen text at the bottom of the screen. The effect size was 0.48, indicating a moderate effect. In this study, placing words near the corresponding portion of the picture helped students understand an explanation of how lightning storms develop, thus yielding what we call the spatial contiguity principle.

Temporal Contiguity Principle

The third principle is that students learn more deeply when corresponding portions of the narration and animation are presented at the same time than when they are separated in time. The theoretical rationale is that learners are better able to make mental connections when corresponding words and pictures are in working memory at the same time. In all eight experimental tests that we conducted, involving explanations of how brakes work, how pumps work, how lightning forms, and how human lungs work, students performed better on tests of problem-solving transfer when animation and narration where presented simultaneously than when they were presented successively (i.e., entire animation before or after entire narration). The median effect size was 1.30, indicating a strong and consistent effect that we call the temporal contiguity effect.

Coherence Principle

The fourth principle is that students learn more deeply from animation and narration when extraneous words, sounds (including music), and video are excluded rather than included. The theoretical rationale is that the learner may attend to the irrelevant material and therefore have less cognitive resource available for building mental connections between relevant portions of the narration and animation. For example, we added extra verbal details or interesting video to an explanation of how lightning works or we added background music and environmental sounds to an explanation of how lightning or brakes work. In five out of five experimental comparisons, students performed better on problem-solving transfer tests when they studied animation and narration without rather than with extraneous words, video, or sounds. The median effect size was 0.90, indicating a strong and consistent effect that we call the coherence effect.

Modality Principle

The fifth principle is that students learn more deeply from animation and narration than from animation and on-screen text. The theoretical rationale is that the learner's visual channel might become overloaded when words and pictures are both presented visually, that is, learners must process the on-screen text and the animation through the eyes, at least initially. Thus, the learner might not have much cognitive capacity left over to build connections between words and pictures. In contrast, when words are presented through the auditory channel (as narration) then the visual channel is less likely to become overloaded, and learners are more likely to be able to build connections between corresponding words and pictures. In six of six experimental comparisons, involving explanations of how lightning forms, how brakes work, and how plants grow, students were better able to transfer what they had learned to new problems when animation was accompanied by spoken words (narration) than by printed words (on-screen text). In all cases, the corresponding animation and words were presented simultaneously. The median effect size was 1.17, indicating a strong and consistent effect, which we call a modality effect.

Redundancy Principle

The redundancy principle is that students learn more deeply from animation and narration than from animation, narration, and on-screen text. It is based on the same theoretical rationale as the modality principle. In two experiments, we compared the problem-solving transfer performance of students who studied an explanation of how lightning forms from animation and corresponding narration versus from animation with corresponding narration and on-screen text. In both studies, receiving less—animation and narration—resulted in better transfer performance than receiving more—animation, narration, and on-screen text. In all cases, the corresponding animation and words were presented simultaneously. The median effect size was 0.77, indicating a moderately strong and consistent effect, which we call the redundancy effect.

Personalization Principle

The final principle is that students learn more deeply from animation and narration when the narration is in conversational rather than formal style. To create the conversational style we added first and second person constructions (i.e., involving "I" and "you") to explanations of lightning formation or plant growth. The theoretical rationale is that students work harder to understand an explanation when they are personally involved in a conversation. In five of five experimental studies, students performed better on transfer tests when words were presented in conversational rather than formal style. The median effect size was 1.55, indicating strong and consistent support for what we call the personalization effect.

Summary

Overall, Table I summarizes seven design principles that flow from our program of research on multimedia learning at UCSB. The effects are most consistent with the cognitive theory of multimedia learning rather than the information delivery theory of multimedia learning. It should be noted that other scholars have found similar effects (see Sweller, 1999, for example) and that user interactivity may be another important principle (see Rieber, 1996).

WHAT IS THE FUTURE OF ANIMATION AS AN AID TO MULTIMEDIA LEARNING?

This review shows that animation has great potential to improve human learning—especially when the goal is to promote deep understanding. However, in order to effectively use animation it is useful to understand how people learn from pictoral and verbal media. Our seven principles are based on a cognitive theory of multimedia learning and are tested in rigorous experimental studies. Yet, our principles should not be taken as rigid procedures to be followed in all situations. Instead, multimedia presentations should be designed in ways that promote the cognitive processes required for meaningful learning, namely selecting, organizing, and integrating as indicated in Fig. 1.

In the new millennium, pictoral forms of teaching are likely to continue to grow as a complement to verbal forms of teaching (Pailliotet and Mosenthal, 2000). Animation is a potentially powerful tool for multimedia designers, but its use should be based on cognitive theory and empirical research. This article provides research-based examples of ways in which animation can be used effectively to promote learner understanding. Yet, animation (and other visual forms of presentation) is not a magical panacea that automatically creates understanding. Indeed, the worldwide web and commercial software are replete with examples of glitzy animations that dazzle the eyes, but it is fair to ask whether or not they promote learner understanding that empowers the mind. Our goal is to develop a cognitive theory of multimedia learning that will guide designers in effectively using animation in multimedia presentations. The future of instructional animation is bright to the extent that its use is guided by cognitive theory and research.

ACKNOWLEDGMENT

This research was supported by a grant from the National Science Foundation entitled "Learning and intelligent systems: Animated pedagogical agents for constructivist learning environments."

REFERENCES

Baddeley, A. (1998). *Human Memory*, Allyn and Bacon, Boston.
Clark, R. E. (1994). Media will never influence learning. *Educ. Technol. Res. Dev.* 42: 21–30.
Dunbar, K. (1993). Concept discovery in a scientific domain. *Cogn. Sci.* 17: 397–434.
Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educ. Technol. Res. Dev.* 42: 7–19.
Mayer, R. E. (1996). Learners as information processors: Legacies and limitations of educational psychology's second metaphor. *Educ. Psychol.* 31: 151–161.
Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educ. Psychol.* 32: 1–19.
Mayer, R. E. (1999). Multimedia aids to problem-solving transfer. *Int. J. Educ. Res.* 31: 661–624.
Mayer, R. E. (in press). *Multimedia Learning*, Cambridge University Press, New York.
Mayer, R. E., and Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *J. Educ. Psychol.* 83: 484–490.
Mayer, R. E., and Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *J. Educ. Psychol.* 84: 444–452.
Mayer, R. E., Heiser, J., and Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *J. Educ. Psychol.* 93: 187–198.
Mayer, R. E., and Moreno, R. (1998). A split-attention affect in multimedia learning: Evidence for dual processing systems in working memory. *J. Educ. Psychol.* 90: 312–320.
Mayer, R. E., Moreno, R., Boire, M., and Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *J. Educ. Psychol.* 91: 638–643.
Mayer, R. E., and Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *J. Educ. Psychol.* 86: 389–401.
Moreno, R., and Mayer, R. E. (1999a). Cognitive principles of multimedia learning: The role of modality and contiguity. *J. Educ. Psychol.* 91: 358–368.
Moreno, R., and Mayer, R. E. (1999b). Multimedia-supported metaphors for meaning making in mathematics. *Cogn. Instr.* 17: 215–248.
Moreno, R., and Mayer, R. E. (2000a). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *J. Educ. Psychol.* 92: 117–125.
Moreno, R., and Mayer, R. E. (2000b). Engaging students in active learning: The case for personalized multimedia messages. *J. Educ. Psychol.* 93: 724–733.

Moreno, R., Mayer, R. E., and Lester, J. C. (2000). Life-like pedagogical agents in constructivist multimedia environments: Cognitive consequences of their interaction. In *Proceedings of ED-MEDIA 2000*, AACE Press, Charlottesville, VA, pp. 741–746.
Pailliotet, A. W., and Mosenthal, P. B. (eds.). (2000). *Reconceptualizing Literacy in the Age of Media, Multimedia, and Hypermedia*, JAI/Ablex, Norwood, NJ.
Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*, Oxford University Press, Oxford, England.
Rieber, L. P. (1990). Animation in computer-based instruction. *Educ. Technol. Res. Dev.* 38: 77–86.
Rieber, L. P. (1996). Animation as feedback in computer-based simulation: Representation matters. *Educ. Technol. Res. Dev.* 44: 5–22.
Ross, S. M. (1994). Delivery trucks or groceries? More food for thought on whether media (will, may, can't) influence learning: Introduction to special issue. *Educ. Technol. Res. Dev.* 42: 5–6.
Salomon, G. (1979/1994). *Interaction of Media, Cognition, and Learning*, Erlbaum, Hillsdale, NJ.
Sweller, J. (1999). *Instructional Design in Technical Areas*, ACER, Camberwell, Australia.
van Merrienboer, J. J. G. (1997). *Training Complex Cognitive Skills*, Educational Technology Publications, Englewood Cliffs, NJ.
White, B. Y. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cogn. Instr.* 10: 1–100.
Wittrock, M. C. (1974). Learning as a generative activity. *Educ. Psychol.* 11: 87–95.