Journal of Experimental Psychology: Applied
2005, Vol. 11, No. 4, 256–265


1076-898X/05/$12.00 DOI: 10.1037/1076-898X.11.4.256

# When Static Media Promote Active Learning: Annotated Illustrations Versus Narrated Animations in Multimedia Instruction

Richard E. Mayer, Mary Hegarty, Sarah Mayer, and Julie Campbell
University of California, Santa Barbara

In 4 experiments, students received a lesson consisting of computer-based animation and narration or a lesson consisting of paper-based static diagrams and text. The lessons used the same words and graphics in the paper-based and computer-based versions to explain the process of lightning formation (Experiment 1), how a toilet tank works (Experiment 2), how ocean waves work (Experiment 3), and how a car's braking system works (Experiment 4). On subsequent retention and transfer tests, the paper group performed significantly better than the computer group on 4 of 8 comparisons, and there was no significant difference on the rest. These results support the static media hypothesis, in which static illustrations with printed text reduce extraneous processing and promote germane processing as compared with narrated animations.

*Keywords:* human–computer interaction, online learning, multimedia learning, textbooks, science education

It is often assumed that animation is the preferred mode for presenting graphics about how things work, but the existing research literature on animation does not offer strong support for this assumption, and research suggesting an advantage of animation has been criticized for methodological problems, such as lack of experimental control (Betrancourt, 2005; Hegarty, Kriz, & Cate, 2003; Tversky, Morrison, & Betrancourt, 2002). The present set of experiments is intended to provide a methodologically sound comparison of the learning outcomes of students who learned scientific explanations with animation and narration and those who learned with static illustrations and text containing equivalent information.

## Rationale

We chose computer-based animation and narration as our dynamic-media treatment because it is the most typical and most recommended mode for presenting animation, and we chose paper-based illustrations and printed text as our static-media treatment because it is the most typical and recommended mode for presenting illustrations. It should be noted that sometimes the most typical or most recommended advice is not the best, but these two treatments also allow us to test useful aspects of a cognitive theory of multimedia learning. The animation and narration materials were constructed by using relevant principles of multimedia design for dynamic media such as presenting the words as spoken text rather than printed text (i.e., the modality principle) and presenting corresponding words and animations at the same time (i.e., the temporal contiguity principle). The illustrations and printed text materials were also constructed by using relevant principles of multimedia design for static media such as presenting the illustrations as a series of frames depicting the major steps in the process (i.e., the segmenting principle) and presenting corresponding words near the diagrams they describe (i.e., the spatial contiguity principle; Mayer, 2001). We constructed the two treatments to be informationally equivalent in the sense that both treatments contained the same words and the same set of core images; however, the animation and narration treatment presented words in spoken form, which may carry more information in terms of the expression of the voice, and presented the pictures in the form of animation containing intervening frames as well as the core frames that were presented in the illustrations and text treatment. For the purposes of this paper, we use the term *paper* (or static media) to refer to paper-based static illustrations and printed text, and we use the term *computer* (or dynamic media) to refer to computer-based animation and narration.

## Practical and Theoretical Goals

Our research has both a practical and a theoretical goal. Our practical goal is to determine whether the general assumptions concerning the value of animation are upheld in a series of scientifically rigorous experiments. Given the additional cost and technical expertise needed to construct and deliver computer-based narrated animations, it is worthwhile to ask whether paper-based static illustrations and printed text can promote learning that is as good or better.

Our theoretical goal is to examine the cognitive processes underlying learning from multimedia. In particular, we tested the *static-media hypothesis,* which states that static media (such as static diagrams and printed text) offer cognitive processing affordances that lead to better learning (as measured by tests of retention and transfer) compared with dynamic media (such as animation and narration). We focus on two measures of learning:

Richard E. Mayer, Mary Hegarty, Sarah Mayer, and Julie Campbell, Department of Psychology, University of California, Santa Barbara.

This research was supported in part by Grants N00014-03-1–0119 and N00014-04-1–0553 from the Office of Naval Research.

Correspondence concerning this article should be addressed to Richard E. Mayer, Department of Psychology, University of California, Santa Barbara, CA 93106. E-mail: mayer@psych.ucsb.edu

retention tests are intended to measure how much was learned and are implemented as recall tasks, and transfer tests are intended to measure how well the learner can apply what was learned to solve new problems and are implemented as essay questions.

## The Static Media Hypothesis and Cognitive Load Theory

The static media hypothesis can be interpreted within the framework of cognitive load theory (Paas, Renkl, & Sweller, 2003; Sweller, 1999, 2005) and the cognitive theory of multimedia learning (Mayer, 2001, 2005), which hold that the attention of learners is limited. Attention can be used for extraneous processing—cognitive processing that does not foster the instructional objective; intrinsic processing—cognitive processing that involves attending to the key material and relations; and germane processing—cognitive processing that involves deeper processing of the key material by mental organization of it into a coherent cognitive representation and integration of it with other representations and prior knowledge. If too much attention is allocated to extraneous processing, there may not be enough remaining attention for intrinsic and germane processing (such as trying to make sense of the presented explanation). On the basis of this analysis, there are two complementary explanations for why learning from static illustrations and printed text could lead to deeper learning than learning from animation and narration: (a) less load from extraneous and intrinsic processing and (b) more germane processing.

The top portion of Table 1 summarizes some possible cognitive advantages of learning from static illustrations and text, on the basis of these assumptions. First, the static illustrations and text treatment enables learners to manage intrinsic processing by allowing them to control the pace and order of presentation. If the pace of presentation begins to overload the learner's cognitive system, the learner can simply slow down his or her rate of reading. If the learner wishes to engage in deeper processing such as mentally connecting corresponding aspects of the words and pictures, the learner is free to look back and forth between corresponding parts of the printed text and illustration. Second, by presenting only frames representing the key steps in the process, the static illustrations and text encourage learners to focus on the most relevant information in the illustrations, thus reducing extraneous processing. Third, by presenting the frames as a series, learners are encouraged to engage in active processing, such as noting the main changes from one frame to the next. This active processing can be characterized as elaboration (Pressley, 1998), self-explanation (Chi, 2000; Roy & Chi, 2005), or mental animation (Hegarty, 2005; Hegarty, Kriz, & Cate, 2003) and corresponds to what Sweller (1999, 2005) calls germane processing and Mayer and Moreno (2003) and Mayer (2005) call essential processing. Accordingly, the increase in germane (or essential processing) should be reflected in better scores on tests of retention and transfer for learners receiving static illustrations and text (i.e., static media) as compared with those receiving animation and narration (i.e., dynamic media).

Table 1
*Cognitive Processes in Learning With Static Illustrations and Text Versus With Animation and Narration*

| Cognitive process |
| --- |
| Static illustrations and text help learners |
| Manage intrinsic processing because learners can control the pace and order of presentation (i.e., learner control effect). |
| Reduce extraneous processing because learners see only frames that distinguish each major step (i.e., signaling effect). |
| Engage in germane processing because learners are encouraged to explain the changes from one frame to the next (active processing effect). |
| Animation and narration help learners |
| Reduce extraneous processing because animation requires less effort to create mental pictorial representation (i.e., effort effect), narration requires less effort to create mental verbal representation (i.e., effort effect), and computer c ontrol requires less effort to make choices during learning (i.e., effort effect). |
| Engage in germane processing because narrated animation creates interest that motivates learners to exert more effort (i.e., interest effect). |

In contrast, the bottom portion of Table 1 lists some possible cognitive advantages of animation and narration as compared with paper-based illustrations and text, which form the basis for the dynamic media hypothesis. First, animations and narrations require less initial cognitive effort to receive the message than do paper-based illustrations and text. That is, the pictures are presented as an animation so the learner does not have to exert cognitive effort to mentally construct a dynamic representation, the words are presented in spoken form so the learner does not have to read, and the pace and order of presentation are predetermined so the learner does not have to exert mental effort in deciding what to do. This case for dynamic media is consistent with the idea that animation is better than static diagrams because it offers a more realistic representation of the to-be-explained process, that is, a representation that is more similar to the real process that it represents. Second, the computer-based animation and narration may be more interesting, entertaining, and motivating than the paper-based illustrations and text, so the learners may exert more effort in making sense of the material—that is, learners may be motivated to engage in germane (or essential) processing. Such a decrease in extraneous processing and increase in germane processing should lead to better scores on tests of retention and transfer for learners receiving animation and narration than those receiving illustrations and text. However, we are somewhat skeptical of the cognitive processing assumptions of the dynamic processing hypothesis because previous research has shown that attempts to increase the interest value of a lesson by adding entertaining features tend to distract learners and lead to poorer learner outcomes (Mayer, 2001).

Thus, proponents of animation might argue that processing of narrated animations requires less cognitive load than does processing of static annotated illustrations, because learners do not have to engage in cognitive processing to animate the graphics when the computer does this for them. However, according to the static media hypothesis, this is the kind of cognitive processing that should be encouraged because it is a form of germane processing that leads to deeper learning. Accordingly, the cognitive cost of animation—which reduces the need for germane processing—is that it increases the need for extraneous processing such as mentally holding previously presented frames in working memory. The transient nature of dynamic media such as narrated animation requires that the learner temporarily holds previously presented material in working memory to later connect it with incoming material. This need for what Mayer and Moreno (2003) call *representational holding* requires cognitive capacity and thereby reduces the amount of cognitive capacity available for essential and generative processing. It is an example of extraneous processing. Trading reduced generative processing for increased extraneous processing is not a good bargain according to the static media hypothesis.

In this study we explore three possible outcomes: the cognitive advantages of static illustrations and text might outweigh the cognitive advantages of animation and narration; the cognitive advantages of animation and narration might outweigh the cognitive advantages of static illustrations and text; or the cognitive advantages of the two treatments in promoting germane processing might be equal, resulting in no differences in transfer test performance. The present experiments allow us to reduce the universe of cognitive explanations concerning our two treatments. This analysis is limited by problems in defining and measuring cognitive capacity (as well as cognitive effort), and by the fact that there is not universal agreement that cognitive capacity is fixed.

The goal of this paper is not to isolate the effects of each feature distinguishing our paper-based static diagram and text treatments from our computer-based animation and narration treatments, because this would require a much larger set of studies, which can be systematically implemented in subsequent research. Instead, our goal in the present set of studies is to compare the learning outcomes of students who learn with our two treatments, to establish a database for testing the static media hypothesis and the dynamic media hypothesis. We addressed this issue in a series of four experiments involving lessons on lightning (see Figure 1), toilet tanks (see Figure 2), ocean waves (see Figure 3), and brakes (see Figure 4), in which we compared the learning outcomes of students who received static lessons using illustrations and printed text with the learning outcomes of students who received dynamic lessons using animation and narration. We chose these four topics because they have been extensively studied in previous research and because we wanted to examine the generality of the comparison of paper-based and computer-based treatments across a wide variety of content areas.

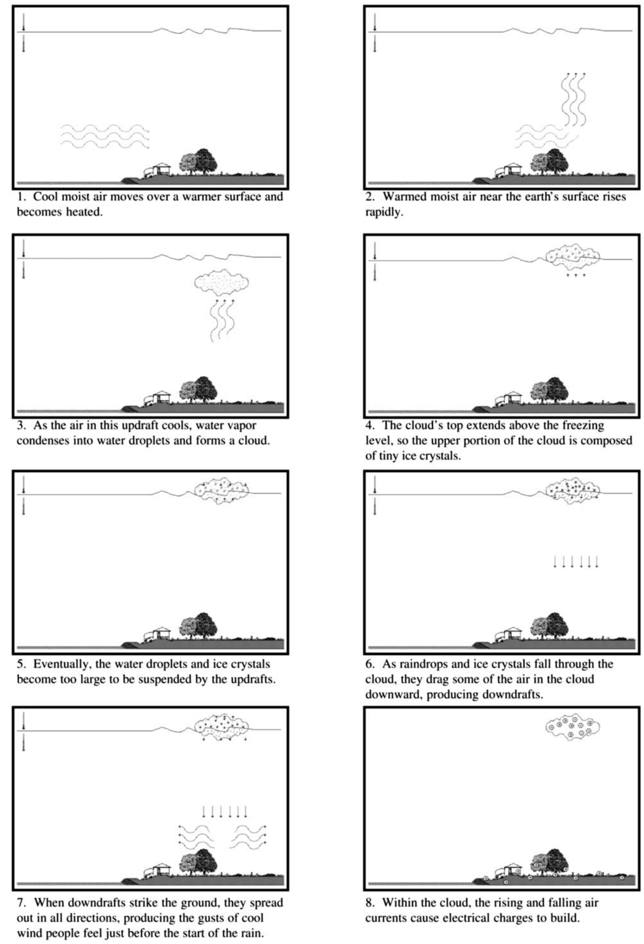



*Figure 1.* Section of paper-based lesson on how lightning works.

## Prior Research on Animation Effects

Previous research comparing the learning outcomes of students learning with text and illustrations versus animation and narration, though sometimes fraught with methodological complications, has generally not produced consistent effects favoring either animation or static media (Hegarty et al., 2003; Hegarty, Narayanan, & Freitas, 2002; Hegarty, Quilici, Narayanan, Holmquist, & Moreno, 1999; Narayanan & Hegarty, 2002; Palmiter & Elkerton, 1993; Palmiter, Elkerton, & Baggett, 1991; Pane, Corbett, & John, 1996; Rieber, 1989; Rieber & Hannafin, 1988; Tversky et al., 2002). In a few studies, an advantage for animation was reported but the animations may have contained more information than the static illustrations or may have involved greater interactivity (Park & Gittelman, 1992; Rieber, 1990, 1991). In the majority of studies, there was no significant difference in the learning outcomes from static and animated media. In a review, Tversky et al. (2002) concluded that animations do not facilitate learning of how complex systems work any better than do static diagrams. However, given the methodological flaws and lack of replicated results noted in many of the studies (Tversky et al., 2002), we sought to provide four clear tests in which computer-based animation and narration was pitted against paper-based diagrams and text. The main advances in this study over some previous work concerns (a) independent variables, that is, the paper-based group and the computer-based group received equivalent information; (b) dependent measures, that is, learning outcomes were assessed with retention and transfer tests; and (c) generality, that is, the comparisons were conducted with similar methodology across four different content areas to determine the generality of the results.

The issue of informational equivalence (Larkin & Simon, 1987) is an important consideration in comparing animation and narration with static illustrations and text. In the studies we report here, we sought to present equivalent information in animation and narration as in illustrations and text. For example, the printed texts in the paper-based treatments have exactly the same words as the narration in the computer-based treatments, and the static diagrams in the paper-based treatments are frames of the animation in the corresponding computer-based treatments. We also allowed ample and identical time for learning from the two types of materials. However, it is difficult (and perhaps impossible) to create narrated animations and annotated illustrations that are exactly information-

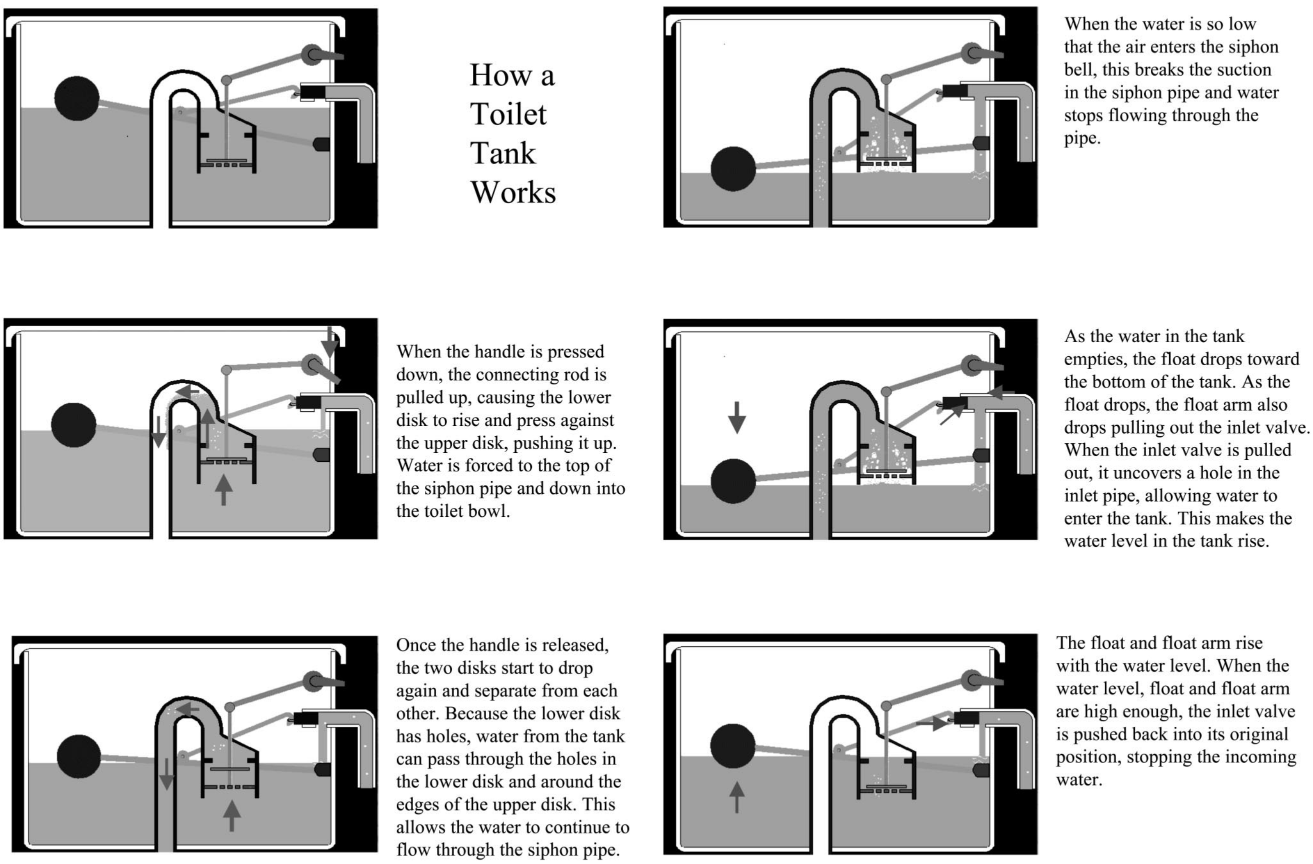



*Figure 2.* Paper-based lesson on how toilet tanks work. (Diagrams were in color.)

ally equivalent. Narration may provide information that is not available in text such as stress and intonation. The static illustrations in the study provided only a small number of the frames of the animations, whereas for complete informational equivalence, they should show every frame. The animations therefore provided information that was not available in static illustrations. For example, an animation might show the exact trajectory of motion of a component, whereas the corresponding static illustrations showed only the beginning and end point of that motion. Thus, we sought to maintain informational equivalence, but on dimensions in which this was not possible, the animation and narration was favored (i.e., added information in terms of the number of frames presented and the tonal quality of the voice).

## Experiment 1

Students learned how lightning develops via computer-based narration and animation or paper-based text and illustrations. Both groups took retention and transfer tests. According to the dynamic media hypothesis the computer group, should outperform the paper group on retention and transfer tests, whereas the opposite pattern is predicted by the static media hypothesis.

### Method

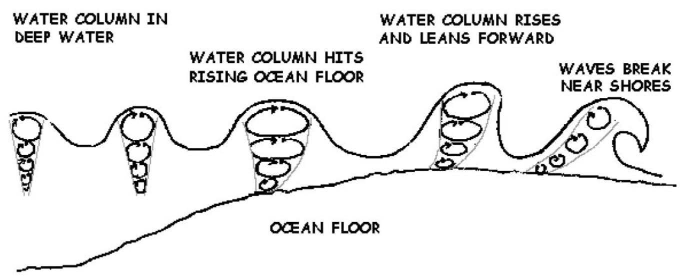
*Participants and design.* The participants were 95 college students recruited from the Psychology Subject Pool at the University of California, Santa Barbara. There were 84 women and 11 men. The mean age was 18.52 years ($SD$ = 1.27), with a range from 18 to 25. The mean SAT score was 1160.91 ($SD$ = 106.05), with a range from 970 to 1350. Fifty-one participants served in the paper group and 44 served in the computer group.

*Materials.* The paper-based materials used in the study included a subject questionnaire, one retention test question, and four transfer test questions. Each question was typed on an 8 1/2 $\times$ 11 in. sheet of paper. The subject questionnaire included questions about the participants' age, gender, and SAT scores. The retention test was a sheet of paper containing the following statement printed at the top "Please write down an explanation of how lightning works." The transfer test consisted of four sheets of paper, each containing one question asking participants about the lesson that they had learned, in a manner that made them apply it to a novel situation. The transfer questions were, "What could you do to decrease the intensity of a lightning storm?" "Suppose you see clouds in the sky, but no lightning. Why not?" "What does air temperature have to do with lightning?" and "What do electrical charges have to do with lightning?"

The paper-based lesson consisted of two facing sheets of paper containing 16 still frames, each depicting a step in the process of lightning, with printed text describing the step under each frame. A page containing the first eight steps is shown in Figure 1. The computer-based lesson contained the same diagrams and the same words as the paper lesson; however, the computer lesson consisted of animated diagrams with concurrent narration. The computer lesson lasted for approximately 140 s and was identical to that used by Mayer and Moreno (1998).

*Apparatus.* The apparatus consisted of five Apple iBook laptop computer systems, with Panasonic headphones.

Near the shore, the ocean floor is not deep enough to allow for the full circular movement of water particles below the surface, so the shallow depth causes the circular orbits to stretch into elliptical orbits. This stretching of the orbits causes the waves to slow down because more time is needed between successive waves. At the same time, when the system of circulating particles (called a water column) hits the rising ocean floor, the crest of the wave is pushed higher above the ocean surface. The top leans forward because it is traveling slightly faster than the bottom, which has slowed down due to colliding with the rising ocean floor. Thus, the shallow depth causes higher crests because the system of circulating water particles (or water column) is pushed upward and is leaning forward. The result is slower, higher-cresting waves.



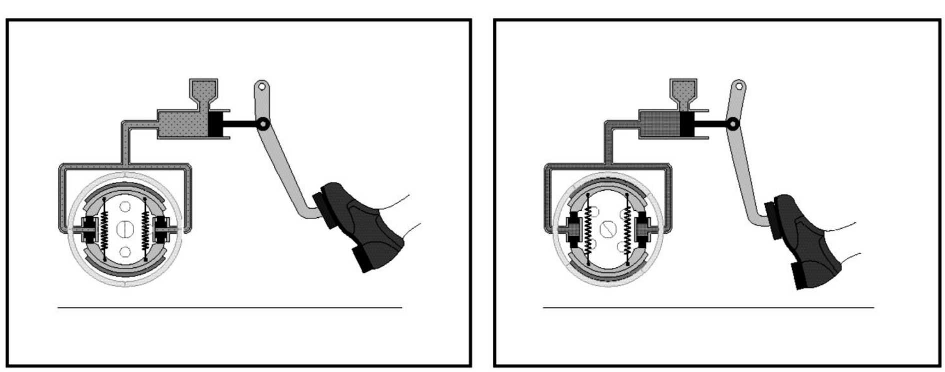

As the crest of a wave rises and tilts forward, a point is reached when its height exceeds its vertical stability and the wave falls into a breaker--seen as white foam--which crashes towards the beach.

*Figure 3.* Section of paper-based lesson on how ocean waves form.

*Procedure.* The participants were tested in groups of 1–5 per session and were randomly assigned to a group. When participants entered the room of the experiment they were each seated in a cubicle. They then filled out a participant questionnaire. Participants in the computer group were seated in front of a computer, and viewed a computer animation and narrative depicting the process of lightning. When the presentation was over, they answered the retention question (with a 4-min time limit) and the four transfer questions (with a limit of 2.5 min each). They received each question sheet for the allotted time limit, and it was removed before the next question sheet was distributed. Participants in the paper group received a paper-based explanation of the process of lightning for 140 s, after filling out the questionnaire. They then answered the same retention and transfer questions, using the same procedure as the computer group.

## Results and Discussion

*Scoring.* For the retention test, a list was produced of 16 major idea units. Participants received one point for each item that they



When the driver steps on the car's brake pedal, a piston moves forward inside the master cylinder.
The piston forces brake fluid out of the master cylinder and through the tubes to the wheel cylinders.
In the wheel cylinders, the increase in fluid pressure makes a smaller set of pistons move.
These smaller pistons activate the brake shoes.
When the brake shoes press against the drum, both the drum and the wheel stop or slow down.

*Figure 4.* Paper-based lesson on how brakes work. (Diagrams were in color.)

Table 2
*Mean Scores, Standard Deviations, and t tests for the Differences Between Paper and Computer Groups on Retention and Transfer Tests: Experiments 1–4*

| | Retention | | | | Transfer | | | |
|---|---|---|---|---|---|---|---|---|
| Experiment and group | *M* | *SD* | *t* | *d* | *M* | *SD* | *t* | *d* |
| Experiment 1 (lightning) | | | | | | | | |
| Paper group ($n$ = 51) | 6.78 | 2.73 | | | 3.65 | 1.58 | | |
| Computer group ($n$ = 44) | 6.24 | 2.66 | | | 2.77 | 1.78 | | |
| | | | 0.96 | 0.20 | | | 3.07* | 0.55 |
| Experiment 2 (toilet tank) | | | | | | | | |
| Paper group ($n$ = 17) | 17.24 | 5.07 | | | 11.82 | 1.91 | | |
| Computer group ($n$ = 14) | 15.43 | 4.54 | | | 9.79 | 2.52 | | |
| | | | 1.04 | 0.36 | | | 2.56* | 1.06 |
| Experiment 3 (waves) | | | | | | | | |
| Paper group ($n$ = 23) | 2.52 | 0.59 | | | 5.09 | 1.56 | | |
| Computer group ($n$ = 17) | 2.00 | 0.86 | | | 4.17 | 1.67 | | |
| | | | 2.26* | 0.88 | | | 1.77 | 0.59 |
| Experiment 4 (brakes) | | | | | | | | |
| Paper group ($n$ = 15) | 3.93 | 1.16 | | | 3.33 | 1.88 | | |
| Computer group ($n$ = 16) | 2.38 | 1.89 | | | 2.50 | 2.13 | | |
| | | | 2.74* | 1.37 | | | 1.15 | 0.44 |

* $p < .05$.

included in their answer to the retention question, regardless of wording. For each transfer question a list of possible correct answers was produced. The participants received one point for each item found on the list that they had included in their answers to the transfer questions. Their final score was then totaled by adding together their scores on each individual transfer question. The Pearson correlation between retention and transfer scores is $r = .375$, $p < .01$.

*Does animation help people remember?* The top-left portion of Table 2 shows the mean scores and standard deviations for each group on the retention test. According to the dynamic media hypothesis, the computer group should score the highest on both retention and transfer because the narrated animation requires less effort to process and creates more interest. According to the static media hypothesis, participants who received the static diagram and text should perform the best because they are encouraged to engage in germane processing and can reduce extraneous processing. In Experiment 1, a *t* test revealed no significant difference between the paper group and the computer group on the retention test, $t(93) = 0.96$, *ns*. The effect size by which the score of paper group exceeds the score of the computer group on the basis of Cohen's *d* was 0.20. The lack of statistical significance renders the results neutral with respect to the hypotheses.

*Does animation help people transfer?* The top-right portion of Table 2 shows the mean scores and standard deviations for each group on the transfer test. A *t* test revealed that the paper group significantly outscored the computer group on the transfer test, $t(93) = 3.07$, $p < .01$, $d = 0.55$. These data are inconsistent with the dynamic media hypothesis and consistent with the static media hypothesis.

## Experiment 2

Students learned how a toilet tank works via computer-based animation and narration or paper-based illustrations and text. Both groups took retention and transfer tests. The methods and predictions were similar to those in Experiment 1.

### Method

*Participants and design.* The participants were 31 college students recruited from the same source as in Experiment 1. There were 26 women and 5 men. The mean age was 18.29 years ($SD = 0.97$), with a range from 17 to 22. The mean SAT score was 1219.60 ($SD = 96.58$), with a range from 1000 to 1360. Seventeen participants served in the paper group and 14 served in the computer group.

*Materials.* The paper-based materials used in the study consisted of a subject questionnaire, an introductory diagram, one retention test question, and four transfer test questions. The materials were essentially the same as in Experiment 1 except as noted in this section. The introductory diagram showed a colored line drawing of the toilet tank with each part labeled with the name used in the lesson. The retention test was a sheet of paper containing the following statement printed at the top: "Describe how a toilet tank works. Imagine that you push down on the handle of the toilet tank. Describe step-by-step what happens to each of the other parts of the tank as it flushes." The transfer test consisted of four questions, each typed on one sheet of paper, which asked the participants to apply the information in the lesson to diagnose possible faults in a malfunctioning toilet tank. The transfer questions were, "Suppose you push down on the handle of the toilet tank but water does not flush into the toilet bowl. Explain all the possible things that could be wrong." "Suppose that after flushing the toilet, you notice that water is continuously running into the toilet tank. Explain all the possible things that could be wrong." "Suppose that after you flush the toilet, water continues to run into the toilet bowl without stopping. Explain all of the possible things that could be wrong." and "What would happen if the float were to break off from the float arm? What would happen if the upper and lower disks were to stick to each other in the siphon bell? What stops the water from flushing out of the tank?"

The paper-based lesson consisted of two facing sheets of paper, with a series of six illustrations and accompanying text explaining the steps in how a toilet tank works. The text was placed to the right of the illustration that it described. Figure 2 shows the two sheets used in the static-media lesson (except that the diagrams were in color).

The computer-based lesson contained the same illustrations and the same words as the paper-based lesson; however it consisted of computer-based animated illustrations with concurrent narration. The computer-based lesson lasted for approximately 78 s and was identical to that used by Hegarty et al. (2003). After they viewed the presentation, learners could click on a button to repeat the presentation either with or without narration.

*Apparatus.* The apparatus consisted of five Apple IBook laptop computer systems, with Panasonic headphones.

*Procedure.* The procedure was the same as in Experiment 1 except as noted in this section. Participants in the computer group were seated in front of a computer and given the introductory diagram for 1 min. Then, they viewed the computer animation and narrative depicting how the toilet tank works. They were allowed 6 min, which enabled them to repeat the presentation. Participants in the paper group received for 6 min a paper-based explanation of how a toilet tank works. After receiving the computer or paper lesson, they answered one retention question (for 6 min) and four transfer questions (for 2.5 min for the first three questions and for 4 min for the last one).

### Results and Discussion

*Scoring.* For the retention test, a list was produced of 33 major idea units. Participants received one point for each item that they included in their answer to the retention question, regardless of wording. For each transfer question a list of possible correct answers was produced. The participants received one point for each item found on the list that they had included in their answers to the transfer questions. Their final score was determined by adding together their scores on each individual transfer question. The Pearson correlation between retention and transfer scores was $r = .219$, *ns*.

*Does animation help people remember?* The left side of the second portion of Table 2 shows the mean scores and standard deviations for each group on the retention test. In Experiment 2, a *t* test revealed no significant difference between the paper group and the computer group on the retention test, $t(29) = 1.04$, *ns*. The effect size, favoring the paper group, was $d = .36$. The lack of statistical significance renders the results neutral with respect to the static media hypothesis.

*Does animation help people transfer?* The right side of the second portion of Table 2 shows the mean scores and standard deviations for each group on the transfer test. A *t* test that the paper group significantly outscored the computer group on the transfer test, $t(29) = 2.56$, $p = < .05$, $d = 1.06$. These data are inconsistent with the dynamic media hypothesis and they support the static media hypothesis.

## Experiment 3

Students learned how ocean waves work via computer-based animation and narration or paper-based illustrations and text. Both groups took retention and transfer tests. The methods and predictions were similar to those in Experiments 1 and 2.

### Method

*Participants and design.* The participants were 40 college students recruited from the same source as in Experiment 1. There were 28 women and 12 men. The mean age was 18.20 ($SD = 0.76$), with a range from 17 to 21. The mean SAT score was 1230.00 ($SD = 87.93$), with a range from 1020 to 1430. Twenty-three participants served in the paper group and 17 served in the computer group.

*Materials and apparatus.* The paper materials included a participant questionnaire, retention sheet, and five transfer sheets. The materials were essentially the same as in Experiments 1 and 2 except as noted in this section. The retention test was a sheet containing the following statement printed at the top: “What causes the formation of ocean waves? Why do they move toward shore? Why do they break at the shore?” The transfer test sheets involved asking the participants about the lesson that they had learned, in a manner that made them apply it to a particular situation. The transfer questions were, “On Tuesday the waves were very high at Sunset Beach. On Wednesday the waves were not very high at Sunset Beach. What caused the decrease in the height of the waves?” “Why do waves rise and curl? Why do they fall?” “If waves keep moving to shore why doesn’t the ocean run out of water?” “What could you do to make sure waves break about 50 feet from shore?” and “The beaches just north of where you live are getting pounded by large breaking waves, however, there are no waves breaking at your local beach. Why not?”

The paper-based lesson consisted of six sheets of paper, with six illustrations and 653 words explaining how ocean waves are created, move across the ocean, and break near the shore. Figure 3 shows the final page from the static-media lesson concerning breaking of ocean waves.

The computer-based lesson contained the same diagrams and the same words as the static-media lesson; however, the dynamic-media lesson consisted of animated diagrams with concurrent narration. The computer-based animation and narration lasted for approximately 7 min and was identical to that used by Mayer and Jackson (2005).

*Apparatus.* The apparatus consisted of five Sony Vaio laptop computers with 15-in. screens and Sony headphones.

*Procedure.* The procedure was the same as in Experiments 1 and 2 except as noted below. Participants in the computer group were seated in front of a computer and viewed a 7-min computer-based animation and narration depicting how ocean waves work. Participants in the paper group received for 7 min a paper-based explanation of how ocean waves work, after filling out the questionnaire. Upon finishing the lesson, they answered one retention question and five transfer questions, with a 2.5-min time limit for each.

### Results and Discussion

*Scoring.* For the retention test a list was produced of 6 major idea units. Participants received one point for each item that they included in their answer to the retention question, regardless of wording. For each transfer question a list of possible correct answers was produced. The participants received one point for each item found on the list that they had included in their answers to the transfer questions. Their final score was determined by adding together their scores on each individual transfer question. The Pearson correlation between retention and transfer scores was $r = .258$, *ns*.

*Does animation help people remember?* The left side of the third portion of Table 2 shows the mean scores and standard deviations for each group on the retention test. In Experiment 3, a *t* test revealed that the paper group significantly outscored the computer group on the retention test, $t(38) = 2.26$, $p < .05$, $d = 0.88$. These data do not support the dynamic media hypothesis and do support the static media hypothesis.

*Does animation help people transfer?* The right side of the third portion of Table 2 shows the mean scores and standard deviations for each group on the transfer test. A *t* test revealed no significant difference between the paper group and the computer group on the transfer test, $t(38) = 1.77$, *ns*. The effect size, favoring the paper group, was $d = 0.59$. These data are inconsistent with the dynamic media hypothesis and are neutral with respect to the static media hypothesis.

## Experiment 4

Students learned how a car's braking systems works via computer-based narration and animation or paper-based text and illustrations. Both groups took retention and transfer tests. The methods and predictions are similar to those in Experiments 1, 2, and 3.

### Method

*Participants and design.* The participants were 31 college students recruited from the same source as in Experiment 1. There were 23 women and 8 men. The mean age was 19.00 ($SD = 1.00$), with a range from 18 to 22. The mean SAT score was 1138.33 ($SD = 114.61$), with a range from 950 to 1330. Fifteen participants served in the paper group and 16 served in the computer group.

*Materials.* The paper-based materials used in the study consisted of a participant questionnaire, one retention test question, and four transfer test questions. The materials were essentially the same as in Experiments 1, 2, and 3 except as noted in this section. The retention test was a sheet of paper containing the following statement printed at the top: "Please write down an explanation of how a car's braking system works. Pretend that you are writing to someone who does not know much about brakes." The transfer test consisted of four sheets of paper, each containing one question asking participants about the lesson that they had learned, in a manner that made them apply it to a novel situation. The transfer questions were, "What could be done to make brakes more reliable, that is, to make sure they do not fail?" "What could be done to make brakes more effective, that is, to reduce the distance needed to bring a car to a stop?" "Suppose you press on the brake pedal in your car but the brakes don't work. What could have gone wrong?" and "What happens when you pump the brakes (i.e., press the pedal and repeatedly and rapidly)?"

The paper lesson consisted of a sheet of paper, with two illustrations and text explaining how a car's braking system works. The illustrations showed the configuration of parts of the braking system in two different states, when the brakes are engaged (one presses down on the brake pedal) and when the brakes are released. Figure 4 shows the paper lesson (except that the diagrams were in color). The computer lesson contained the same diagrams and the same words as the paper lesson; however, the computer lesson consisted of animated diagrams with concurrent narration. The computer-based animation and narration lasted for approximately 50 s and was identical to that used by Mayer and Anderson (1992).

*Apparatus.* The apparatus consisted of five Apple iBook laptop computer systems, with Panasonic headphones.

*Procedure.* The procedure was the same as in Experiments 1, 2, and 3 except as noted in this section. Participants in the computer group were seated in front of a computer and viewed a computer animation and narrative depicting how brakes work. Participants in the paper group received a paper-based explanation of how brakes work, for 1 min, after filling out the questionnaire. When the computer or paper lesson was over, they answered the retention question and the four transfer questions, for 2.5 min per question.

### Results and Discussion

*Scoring.* For the retention test a list was produced of eight major idea units. Participants received one point for each item that they included in their answer to the retention question, regardless of wording. For each transfer question a list of possible correct answers was produced. The participants received one point for each item found on the list that they had included in their answers to the transfer questions. Their final score was then totaled by adding together their scores on each individual transfer question. The Pearson correlation between retention and transfer score was $r = .523$, $p < .01$.

*Does animation help people remember?* The bottom-left portion of Table 2 shows the mean scores and standard deviations for each group on the retention test. According to the dynamic media hypothesis, the computer group should score the highest on both retention and transfer, and according to the static media hypothesis, participants who received the static diagrams and text should perform the best. In Experiment 4, a $t$ test revealed that the paper group significantly outscored the computer group on the retention test, $t(29) = 2.74$, $p < .05$, $d = 1.07$. These data support the static media hypothesis and are inconsistent with the dynamic media hypothesis.

*Does animation help people transfer?* The bottom-right portion of Table 2 shows the mean scores and standard deviations for each group on the transfer test. A $t$ test revealed no significant difference between the paper group and the computer group on the transfer test, $t(29) = 1.15$, *ns*. The effect size, favoring the paper group, was $d = .44$. The lack of statistical significance renders the difference neutral with respect to the static media hypothesis.

The results of Experiment 4 should be interpreted in light of the fact that the paper group had up to 60 s to read the lesson, whereas the computer group received a presentation that was 50 s long. The rationale for giving participants in the paper group a few more seconds is that it might take them a few seconds to get oriented before beginning to read the lesson, whereas the 50-s time limit for the computer lesson did not begin until the participant pressed a button.

## General Discussion

### Theoretical Implications

Does a dynamic medium (involving computer-based animation and narration) help learning compared with a static medium (involving paper-based static illustrations and printed text)? The dynamic media hypothesis predicts that the computer group will outperform the paper group because students in the computer group expend less effort in perceiving the presentation and engage in deeper cognitive processing of the presentation than people in the paper group. In the studies we conducted, students who learned from computer-based narrated animation (which can be called *dynamic media*) did not score significantly better on posttests than did students who learned from paper-based illustrations and printed text (which we can be called *static media*). In eight comparisons, across four different studies, the computer group never scored higher than the paper group.

Does a dynamic medium (involving computer-based narrated animation) result in less learning compared with a static medium (involving paper-based static illustrations and printed words)? In four of eight comparisons, across four different studies, the paper group scored significantly higher than the computer group. In the remaining four comparisons, the difference failed to reach statistical significance. On the basis of a binomial probability test the probability is less than .001 that the paper group would outscore the computer group by chance on eight of eight tests. Overall, the effect sizes—all favoring the paper group—were 1.37, 1.06, .88, .59, .55, .44, .36, and .22. The median effect size favoring the paper group was .62 for retention and .57 for transfer. Effect size analyses are a particularly helpful tool, especially in light of the low sample size in some studies and the variety of experimental contexts. They help to show concisely that there was support for the static media hypothesis—that is, some evidence that people engage in less extraneous processing and therefore are able to engage in deeper cognitive processing when they learn from static illustrations and text rather than the dynamic animations and commentaries.

Although many studies (e.g., as reviewed by Tversky et al., 2002) have found no difference between static and animated media, an important novel contribution of this series of experiments is that they are the first replicated empirical demonstrations that static presentations containing illustrations and printed text can actually be superior to dynamic presentations containing narrated animation in terms of learning outcomes. The advantage of the static-media presentation is that it can reduce extraneous processing (such as attending to unimportant movement in the animation and having to hold animation frames in working memory to mentally link one to the next), and it can promote generative processing (such as mentally animating or self-explaining the key changes from one static frame to another). However, it should be noted that in four of eight comparisons there was no significant difference between the paper and computer groups. A limitation of this study is that there is insufficient power to detect small differences that might underlie the phenomenon. In addition, unequal cell sizes can affect the results.

### *Practical Implications*

On a practical level, these results suggest caution in the use of animation to depict the operation of physical and mechanical systems. Trying to make a presentation less effortful and more interesting through the use of computer-generated animation led to worse test performance on four of eight comparisons and no significant difference on the others. Animation may be entertaining, but these experiments offer no reason to conclude that animation inherently provides more educational value than static diagrams. Instead, a well-designed series of still frames can be as good or better than animation in promoting learning. Although some educators may believe that technological innovations such as computer-based animations are the wave of the future, this study suggests that it may be premature to toss out the older technology embodied in textbooks. It should be noted that the studies reported in this article are based on a highly selected population (i.e., college students at a selective university), so the results might not generalize to a population that includes lower ability or lower literacy individuals.

### *Future Directions*

What can account for the surprising finding that illustrations and text can sometimes lead to better learning outcomes than animation and narration? First, the paper treatment involves simultaneous presentation of the frames whereas the computer presentation involves successive presentation of the frames. Second, the paper treatment is learner controlled because the learner can determine the pacing and order of the presentation simply through eye movements, and the computer treatment is instructor controlled because the narrated animation is presented at a fixed pace and order (see Mayer & Chandler, 2001; Zacks & Tversky, 2003). Third, in the paper treatment the presentation is segmented into meaningful units showing crucial states of the system, whereas in the computer treatment the presentation is a continuous unit (see Hegarty, 1992). Fourth, words are printed in the paper treatment and words are spoken in the computer treatment. Fifth, the paper group received the pictorial material on one or two facing sheets of paper, whereas the computer group received the pictorial material on a computer screen, so although it is implausible, we cannot rule out the possibility that the effects are due to differences in the delivery devices used to present the information. Future research is needed to disentangle which features of the paper and computer treatments contribute to differences in test performance. Other factors worthy of future study include the selection of frames for the static presentation, the role of presentation speed for the animation, and the placement of cuts in the animation.

## Conclusion

Overall, the main contribution of this set of experiments is that there is no support for the dynamic media hypothesis and there is support for the static media hypothesis when informationally equivalent paper and computer treatments are compared on retention and transfer measures. It is important to note that the findings are consistent across four different content areas.

This study should not be interpreted to mean that animations are ineffective in all situations. Although animations did not improve test performance in this set of experiments, there may be situations in which animations improve understanding. For example, learners have limitations in spatial ability (Hegarty & Sims, 1994), so that with more complex systems, they may not always be able to mentally animate how a system works from a series of static diagrams. Animations may also be more effective when they are used to visualize processes that are not visible in the real world, such as the movement of air around pressure systems in meteorology, or showing how a computer sort algorithm works (Narayanan & Hegarty, 2002). ChanLin (1998) has shown that animation may have advantages over static illustrations for certain kinds of information and certain kinds of learners.

Overall, this research should not be taken to controvert the value of animation as an instructional aid to learning. Instead, this research suggests that when computer-based animations are used in instruction, learners may need some assistance in how to process these animations. This research suggests that animations could be constructed in ways that tap the positive features of static illustrations. For example, learners can be given control over the pace and order of animations by being allowed to use slider bars and pause buttons; learners can be guided to attend to the key steps in an animation by presentation of the animation in meaningful segments in which the next segment is initiated by a learner action such as clicking a “continue” button; and learners can be encouraged to engage in active processing through activities such as generating explanations or answering questions during learning. Subsequent research is needed to determine how best to incorporate devices for learner control, segmenting, and active processing in animations while still retaining their positive features.

In conclusion, this research makes the important point that all graphics are not equally effective. More research is needed to pinpoint the characteristics of effective graphics. On the basis of the static media hypothesis, we recommend showing a series of static frames that clearly depict the possible states of each part in the system so that learners can infer the change from one to the next. Although there is intuitive appeal to the idea of using animated displays to show dynamic processes, the current studies found computer-based narrated animations to be inferior to paper-based annotated illustrations in four of eight comparisons.

## References


Betrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 287–296). New York: Cambridge University Press.

ChanLin, L.-J. (1998). Animation to teach students of different knowledge levels. *Journal of Instructional Psychology, 25,* 166–175.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161–238). Mahwah, NJ: Erlbaum.

Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1084–1102.

Hegarty, M. (2005). Multimedia learning about physical systems. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 447–465). New York: Cambridge University Press.

Hegarty, M., Kriz, S., & Cate, C. (2003). The role of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction, 21,* 325–360.

Hegarty, M., Narayanan, N. H., & Freitas, P. (2002). Understanding machines from multimedia and hypermedia presentations. In J. Otero, A. C. Graesser, & J. Leon (Eds.), *The psychology of science text comprehension* (pp. 357–384). Mahwah, NJ: Erlbaum.

Hegarty, M., Quilici, J., Narayanan, N. H., Holmquist, S., & Moreno, R. (1999). Multimedia instruction: Lessons from evaluation of a theory-based design. *Journal of Educational Multimedia and Hypermedia, 8,* 119–150.

Hegarty, M., & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory & Cognition, 22,* 411–430.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65–100.

Mayer, R. E. (2001). *Multimedia learning.* New York: Cambridge University Press.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.

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

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93,* 390–397.

Mayer, R. E., & Jackson, J. (2005). Concise science instruction: When quantitative details hurt qualitative understanding. *Journal of Experimental Psychology: Applied, 11,* 13–18.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 90,* 312–320.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38,* 43–52.

Narayanan, N. H., & Hegarty, M. (2002). Multimedia design for communication of dynamic information. *International Journal of Human–Computer Studies, 57,* 279–315.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38,* 1–4.

Palmiter, S. L., & Elkerton, J. (1993). Animated demonstrations for learning procedural computer-based tasks. *Human–Computer Interaction, 8,* 193–196.

Palmiter, S. L., Elkerton, J., & Baggett, P. (1991). Animated demonstrations vs. written instructions for learning procedural tasks: A preliminary investigation. *International Journal of Man–Machine Studies, 34,* 687–701.

Pane, J. F., Corbett, A. T., & John, B. E. (1996). Assessing dynamics in computer-based instruction. In M. J. Tauber (Ed.), *Proceedings of the ACM conference on human factors in computing systems* (pp. 797–804). New York: ACM Press.

Park, O. C., & Gittelman, S. S. (1992). Selective use of animation and feedback in computer-based instruction. *Educational Technology, Research and Development, 40,* 125–167.

Pressley, M. (1998). *Reading instruction that works: The case for balanced teaching.* New York: Guilford Press.

Rieber, L. P. (1989). The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. *Journal of Educational Computing Research, 5,* 431–444.

Rieber, L. P. (1990). Using computer animated graphics with science instruction with children. *Journal of Educational Psychology, 83,* 135–140.

Rieber, L. P. (1991). Animation, incidental learning, and continuing education. *Journal of Educational Psychology, 83,* 318–328.

Rieber, L. P., & Hannafin, M. J. (1988). Effects of textual and animated orienting activities and practice on learning from computer-based instruction. *Computers in the Schools, 5,* 77–89.

Roy, M., & Chi, M. T. H. (2005). The self-explanation principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 271–286). New York: Cambridge University Press.

Sweller, J. (1999). *Instructional design in technical areas.* Camberwell, Australia: ACER Press.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human–Computer Studies, 57,* 247–262.

Zacks, J. M., & Tversky, B. (2003). Structuring information interfaces for procedural learning. *Journal of Experimental Psychology: Applied, 9,* 88–100.


Received June 24, 2005
Revision received August 22, 2005
Accepted August 26, 2005 ■