Journal of Educational Psychology
1998, Vol. 90, No. 2, 312-320

0022-0663/98/$3.00

A Split-Attention Effect in Multimedia Learning: Evidence for Dual Processing Systems in Working Memory

Richard E. Mayer and Roxana Moreno
University of California, Santa Barbara

Students viewed a computer-generated animation depicting the process of lightning formation (Experiment 1) or the operation of a car's braking system (Experiment 2). In each experiment, students received either concurrent narration describing the major steps (Group AN) or concurrent on-screen text involving the same words and presentation timing (Group AT). Across both experiments, students in Group AN outperformed students in Group AT in recalling the steps in the process on a retention test, in finding named elements in an illustration on a matching test, and in generating correct solutions to problems on a transfer test. Multimedia learners can integrate words and pictures more easily when the words are presented auditorily rather than visually. This split-attention effect is consistent with a dual-processing model of working memory consisting of separate visual and auditory channels.

In multimedia learning, information is presented to learners in two or more formats, such as in words and in pictures (Mayer, 1997). For example, Figure 1 provides selected frames from a short animation depicting a cause-and-effect explanation of how lightning forms along with corresponding on-screen text which provides the explanation in words. To design effective multimedia presentations, it is useful to understand how learners mentally integrate words and pictures. The purpose of this study is to contribute to multimedia learning theory by testing a dual-processing theory of working memory.

This can be done by comparing the learning outcomes of students who view the lightning animation along with corresponding on-screen text (Group AT) and those who view the lightning animation along with concurrent auditory narration using the same words as the text (Group AN). To assess students' understanding of the material, we asked them to write explanations of how lightning forms (retention test), to give names for parts of an illustration (matching test), and to apply what they learned to solve new problems (transfer test). Example items are presented in Table 1.

Dual-Processing Theory of Working Memory

Our research was designed to test a straightforward prediction of a dual-processing theory of working memory, as summarized in Figure 2. The primary assumptions of dual-processing theory are as follows: (a) Working memory includes an auditory working memory and a visual working memory, which are analogous to the phonological loop and visuospatial sketch pad, respectively, in Baddeley's (1986, 1992) theory of working memory; (b) each working memory store has a limited capacity, consistent with Sweller's (1988, 1989; Chandler & Sweller, 1992; Sweller, Chandler, Tierney, & Cooper, 1990) cognitive load theory; (c) meaningful learning occurs when a learner retains relevant information in each store, organizes the information in each store into a coherent representation, and makes connections between corresponding representations in each store, analogous to the cognitive processes of selecting, organizing, and integrating in Mayer's (1997; Mayer, Steinhoff, Bower, & Mars, 1995) generative theory of multimedia learning; and (d) connections can be made only if corresponding pictorial and verbal information is in working memory at the same time, corresponding to referential connections in Paivio's (1986; Clark & Paivio, 1991) dual-coding theory.

According to the dual-processing theory, visually presented information is processed—at least initially—in visual working memory whereas auditorily presented information is processed—at least initially—in auditory working memory. For example, in reading text, the words may initially be represented in visual working memory and then be translated into sounds in auditory working memory. In the AN treatment, students represent the animation in visual working memory (such as the image of negative signs moving to the bottom of a cloud) and represent the corresponding narration in auditory working memory (such as the statement "negative ions fall to the bottom of the cloud"). Because they can hold corresponding pictorial and verbal representations in working memory at the same time, students in Group AN are better able to build referential connections between them, such as seeing that the image of negative signs moving to the bottom of the clouds corresponds to the words describing the fall of negative ions.

In the AT treatment, students try to represent both the animation and the on-screen text in visual working memory. Although some of the visually represented text eventually

Roxana Moreno created the multimedia materials used in Experiment 1; Matt Mendrala created the multimedia materials used in Experiment 2.

Correspondence concerning this article should be addressed to Richard E. Mayer or Roxana Moreno, Department of Psychology, University of California, Santa Barbara, California 93106. Electronic mail may be sent to mayer@psych.ucsb.edu or to moreno@psych.ucsb.edu.

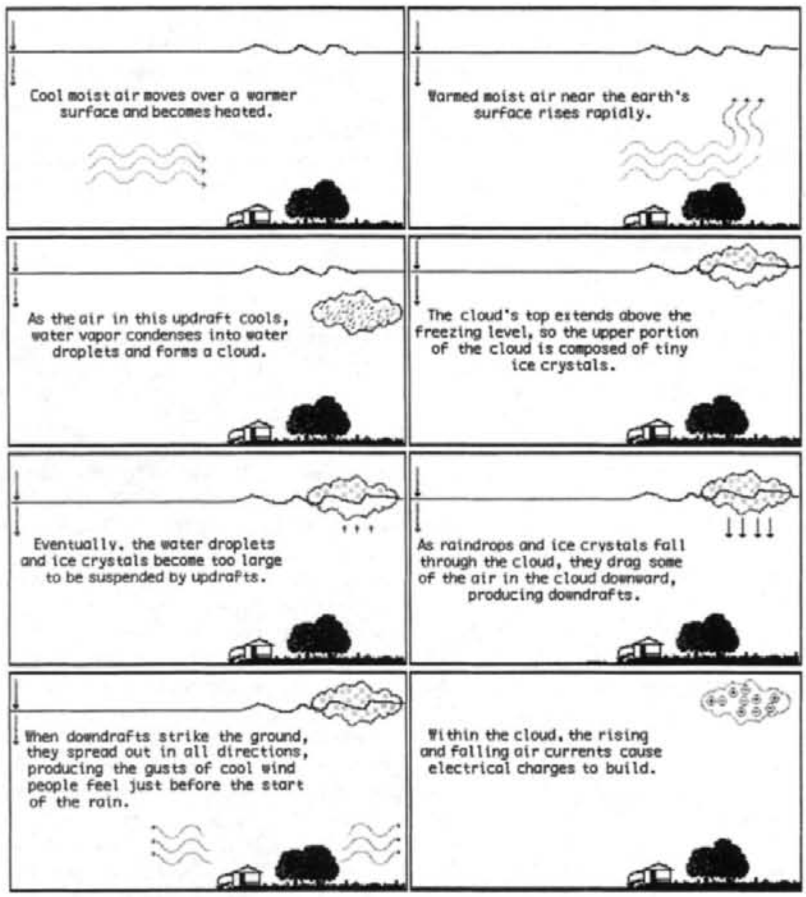

Figure 1. Selected frames from a multimedia lesson on the formation of lightning.

may be translated into an acoustic modality for auditory working memory, visual working memory is likely to become overloaded. Students in Group AT must process all incoming information—at least initially—through their visual working memory. Given the limited resources students have for visual information processing, using a visual modality to present both pictorial and verbal information can create an overload situation for the learner. If students pay full attention to on-line text, they may miss some of the

Table 1
Three Kinds of Questions About the Formation of Lightning

Test	Instruction
Retention	Write an explanation of how lightning works.
Matching	Using the frames shown in Figure 1, circle the cool moist air and write C next to it, circle the warmer surface and write W next to it, circle the updraft and write U next to it, circle the freezing level and write F next to it, circle the downdraft and write D next to it, circle the cool gusts of cool wind and write G next to it, circle the stepped leader and write S next to it, and circle the return stroke and write R next to it.
Transfer	What could you do to decrease the intensity of lightning?
	Suppose you see clouds in the sky but no lightning. Why not?
	What causes lightning?
	What does air temperature have to do with lightning?

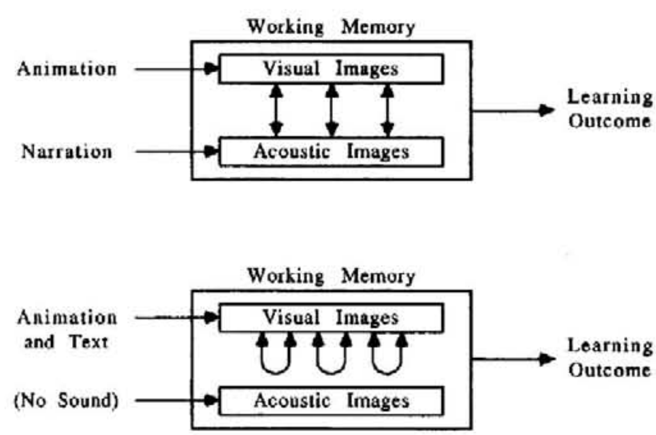

Figure 2. A dual-processing model of multimedia learning. For Group AN, shown in the top panel, the incoming animation and narration initially are held in different working memory spaces. For Group AT, shown in the bottom panel, the incoming animation and text initially are held in the same memory space.

crucial images in the animation, but if they pay full attention to the animation, they may miss some of the on-line text. Because they may not be able to hold corresponding pictorial and verbal representations in working memory at the same time, students in Group AT are less able to build connections between these representations.

Dual-processing theory predicts a split-attention effect in which students in Group AT perform more poorly than students in Group AN on retaining the steps in the cause-and-effect chain (retention test), on being able match pictures and names of parts (matching test), and on being able to use what they have learned to solve problems (transfer test). The prediction of a split-attention effect for retention is based on the idea that AT students may not have encoded as much of the verbal material as AN students. The prediction of a split-attention effect for matching is based on the idea that AT students may not have been able to build as many referential connections between corresponding pictorial and verbal information as AN students did. The prediction of a split-attention effect for transfer is based on the idea that AT students may not have been able to construct a coherent mental model of the system as well as AN students could.

In contrast to the predictions of the dual-processing hypothesis, the information–equivalency hypothesis predicts no differences between the AN and AT groups on any of the tests because identical information was presented to both groups.

Advances in Multimedia Learning Theory

In a study involving static diagrams for geometry problems, Mousavi, Low, and Sweller (1995) found that students learned better when an auditory narration was presented simultaneously with corresponding diagrams (different-modality presentation) than when printed text was presented simultaneously with corresponding diagrams (same-modality presentation). The present study initiated the next phase in research on split-attention effects by examining how students integrate animation with concurrent text that is presented visually or auditorily. In particular, our research is the first to test whether students learn better when an auditory narration is presented along with a corresponding animation (different-modality presentation) than when printed text is presented with a corresponding animation (same-modality presentation). Thus, the present study extends Mousavi et al.'s pioneering examination of a split-attention effect in three ways: (a) by examining a split-attention effect in a computer-based multimedia environment rather than a paper-based environment; (b) by using multiple dependent measures, including transfer and matching; and (c) by using cause-and-effect explanations as the target material.

Experiment 1

In this study, students viewed an animation depicting the process of lightning with concurrent narration (Group AN) or with concurrent on-screen text (Group AT). According to the dual-processing hypothesis, Group AN should perform better than Group AT on recalling relevant steps in the process of lightning formation (retention test), on choosing the correct names for elements in an illustration (matching test), and on generating answers to problems that require applying learning to new situations (transfer test). In contrast, according to the information–equivalency hypothesis, Group AN and Group AT should not differ on any of these tests because both groups received the same information presented for the same length of time.

Method

Participants and design. The participants were 78 college students recruited from the psychology subject pool at the University of California, Santa Barbara. All participants indicated that they lacked experience in meteorology. Forty participants served in the AN group, and 38 participants served in the AT group.

Materials and apparatus. For each participant, the paper-and-pencil materials consisted of a participant questionnaire, a retention test, a matching test, and a four-page transfer test, with each typed on 8.5 × 11 in. (21.59 × 27.94 cm) sheets of paper. The participant questionnaire solicited information concerning the participant's Scholastic Assessment Test (SAT) scores, gender, and meteorology knowledge. Meteorology knowledge was assessed by using a six-item knowledge checklist and a five-item self-rating. The self-assessment asked the participants to rate their knowledge of weather by placing a check mark next to: *very little,* between *very much* and *average, average,* between *average* and *very much,* or *very much.* The checklist consisted of instructions to "please place a check mark next to the items that apply to you" followed by a list of six items: "I regularly read the weather maps in the newspaper," "I know what a cold front is," "I can distinguish between cumulous and nimbus clouds," "I know what a low pressure system is," "I can explain what makes the wind blow," "I know what this symbol means: [symbol for cold front]," and "I know what this symbol means: [symbol for warm front]."

The retention test contained the following instructions at the top of the sheet: "Please write down an explanation of how lightning works."

The matching test presented four frames from the animation along with the following instructions:

> Circle cool moist air and write C next to it. Circle the warmer surface and write W next to it. Circle the updraft and write U next to it. Circle the freezing level and write F next to it. Circle the downdraft and write D next to it. Circle the gusts of cool wind and write G next to it. Circle the stepped leader and write S next to it. Circle the return stroke and write R next to it.

The transfer test consisted of the following four questions, each typed on a separate sheet: "What could you do to decrease the intensity of lightning?" "Suppose you see clouds in the sky, but no lightning. Why not?" "What does air temperature have to do with lightning?" and "What causes lightning?"

The computerized materials consisted of two computer programs for multimedia presentations on how the lightning process works. Both programs generated an identical 140-s animation depicting air moving from the ocean to the land, water vapor condensing to form a cloud, the rising of the cloud beyond the freezing level, the formation of crystals in the cloud, the movement of updrafts and downdrafts, the building of electrical charges within the cloud, the division of positive and negative charges, the traveling of a negative stepped leader from the cloud to the ground, the traveling of a positive stepped leader from the ground to the cloud, the negative charges following the path to the ground, the meeting of the negative leader with the positive leader, and the positive charges following the path towards the cloud. The AN version also included concurrent narration describing each of the major events in words spoken at a slow rate by a male voice. The AT version included concurrent text presented on the screen using the same words and timing as the narration. The computerized materials were adapted from text and illustrations used in previous studies (Mayer, Bove, Bryman, Mars, & Tapangco, 1996; Mayer et al., 1995). Figure 1 presents selected frames from the animation along with on-screen text from the AT version. The multimedia animations were developed using Director 4.0 and Soundedit 16 (Macromedia, 1994).

The apparatus consisted of five Macintosh IIci computer systems, which included 14-in. monitors and Sony headphones.

Procedure. Participants were tested in groups of 1 to 5 per session. Each participant was randomly assigned to a treatment group (either AN or AT) and was seated at an individual cubicle in front of a computer.

First, participants completed the participant questionnaire working at their own rates. Second, the experimenter presented oral instructions stating that the computer would show an animation of how the process of lightning works and that when the computer was finished, the experimenter would have some questions for the participants to answer. Participants in the AN treatment were told to put on headphones, and all participants were told to press the space bar to begin the presentation. Third, after pressing the space bar, the animation with narration was presented once to the participants in the AN group and the animation with text was presented once to the participants in the AT group. Fourth, when the presentation was finished, the experimenter presented oral instructions for the test, stating that there would be a series of question sheets and that for each the participant should keep working until told to stop. Fifth, the retention test was distributed along with instructions to write down an explanation of how lightning works. After 6 min, the test sheet was collected. Then the transfer test sheets were presented one at a time for 3 min each, with each sheet collected by the experimenter before the subsequent sheet was handed out. Then, the matching test was presented and collected after 3 min.

Scoring. A scorer who was not aware of the treatment condition of each participant determined the retention score, matching score, and transfer score for each participant. A retention score was computed for each participant by counting the number of major idea units (out of eight possible) that the participant produced on the retention test. One point was given for correctly stating each of the following eight idea units regardless of wording: (a) air rises, (b) water condenses, (c) water and crystals fall, (d) wind is dragged downward, (e) negative charges fall to the bottom of the cloud, (f) the leaders meet, (g) negative charges rush down, (h) positive charges rush up.

A matching score was computed for each participant by counting the number of correctly labeled elements (out of eight possible) on the matching test. Participants received 1 point for each part that was circled and labeled with the appropriate letter.

A transfer score was computed for each participant by counting the number of acceptable answers (out of 12 possible) that the participant produced across the four transfer problems. For example, acceptable answers for the first question about decreasing lightning intensity included removing positive ions from the ground; acceptable answers for the second question about the reason for the presence clouds without lightning included stating that the tops of the clouds might not be high enough to freeze; acceptable answers for the third question about temperature's relation to lightning included stating that the air must be cooler than the ground; acceptable answers for the fourth question about the causes of lightning included the difference in electrical charges in the cloud. Participants received no more than three points per problem.

Because previous studies have demonstrated that some instructional effects were stronger for low-experience learners than for high-experience learners (Mayer & Gallini, 1990; Mayer & Sims, 1994), we included only low-experience students in our study. We computed an experience score by tallying the number of domain-related activities that the participant checked on the activity checklist and adding that number to the level of experience the participant checked on the five-level self-assessment (with *very little* counted as 0 points *less than average* as 1, *average* as 2, more than *average* as 3, and *very much* as 4). We eliminated the data for any student who scored above 5 and replaced it with the data of a new student. Using this procedure, 16 students were replaced in Experiment 1.

Results and Discussion

Split-attention effect on verbal recall. According to the dual-processing hypothesis, students should remember more of the verbal material when it is presented as narration than when it is presented as text. The retention test bars in the left portion of Figure 3 show the proportion of the eight idea units correctly recalled by students in the AN group and the AT group. An analysis of variance (ANOVA) was conducted with group (AN vs. AT) as the between-subjects factor and retention score as the dependent measure. As can be seen in the retention test bars of Figure 3, there was a split-attention effect in which AN students tended to recall more relevant idea units than did AT students, $F(1, 76) = 15.987$, $MSE = 2.187$, $p < .001$. These results are consistent with the predictions of the dual-processing hypothesis.

Split-attention effect on visual–verbal matching. According to the dual-processing hypothesis, students should perform better on the matching test when verbal material is presented as narration than when it is presented as text. The matching test bars in the middle section of Figure 3 show the proportion correct on the eight-item matching test by students in the AN group and the AT group. An ANOVA was

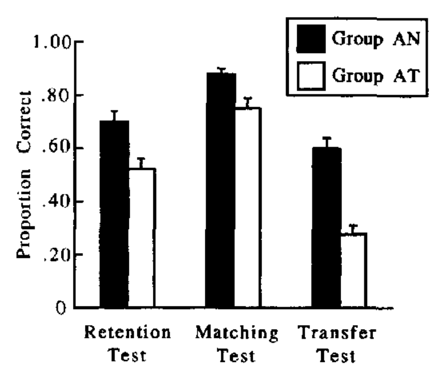

Figure 3. Proportion correct response on retention, matching, and transfer tests by two groups—Experiment 1. On retention, $M = .69$ ($SD = .18$) for Group AN and $M = .52$ ($SD = .19$) for Group AT; on matching, $M = .87$ ($SD = .16$) for Group AN and $M = .77$ ($SD = .22$) for Group AT; on transfer, $M = .60$ ($SD = .24$) for Group AN and $M = .28$ ($SD = .19$) for Group AT. AN = Group receiving concurrent auditory narration; AT = group receiving on-screen text.

conducted with group (AN vs. AT) as the between-subjects factor and matching score as the dependent measure. As can be seen in Figure 3, there was a split-attention effect in which AN students tended to correctly match more items than did AT students, $F(1, 76) = 7.805$, $MSE = 2.380$, $p < .01$. These results are consistent with the predictions of the dual-processing hypothesis.

Split-attention effect on problem-solving transfer. According to the dual-processing hypothesis, students should generate more problem-solving solutions when verbal material is presented as narration than when presented as text. The transfer test bars in the right portion of Figure 3 show the proportion of correct solutions on the transfer test by students in the AN group and the AT group (with the maximum score set at 12). An ANOVA was conducted with group (AN vs. AT) as the between-subjects factor and problem-solving score as the dependent measure. As can be seen in the transfer test bars in the right portion of Figure 3, there was a split-attention effect in which AN students tended to generate more solutions than did AT students, $F(1, 76) = 44.797$, $MSE = 1.683$, $p < .001$. These results are consistent with the predictions of the dual-processing hypothesis.

Overall, split-attention effects were obtained on the retention, matching, and transfer tests, yielding consistent evidence for the dual-processing hypothesis and against the information-equivalency hypothesis. On each test, the effect sizes were substantial: 0.89 for the retention test (based on $SD = 0.19$), 0.55 for the matching test (based on $SD = 0.22$), and 1.75 for the transfer test (based on $SD = 0.19$).

Experiment 2

In the second study, students viewed an animation depicting how a car's braking system works with concurrent narration (Group AN) or with concurrent on-screen text (Group AT). According to the dual-processing hypothesis, Group AN should perform better than Group AT on recalling relevant steps in the process (retention test), on choosing the correct names for parts in an illustration (matching test), and on generating solutions to questions about troubleshooting, redesigning, or explaining braking systems (transfer test). According to the information-equivalency hypothesis, Group AN and Group AT should not differ on any of these tests because both groups received the same information presented for the same length of time.

Method

Participants and design. The participants were 68 college students recruited from the psychology subject pool at the University of California, Santa Barbara. All participants indicated that they had low knowledge of car mechanics. Thirty-four participants served in the AN group, and 34 participants served in the AT group.

Materials and apparatus. The paper-and-pencil materials consisted of a participant questionnaire, a retention test, a matching test, and a four-page transfer test, with each typed on 8.5 × 11 in. (21.59 × 27.94 cm) sheets of paper. The participant questionnaire solicited information concerning the participant's SAT scores, gender, and mechanical experience. Mechanical experience was assessed using an six-item activity checklist and a five-item self-rating. The checklist consisted of instructions to "please place a check mark next to the things you have done" followed by list of six items: "I have a driver's license," "I have put air into a tire on a car," "I have changed a tire on a car," "I have changed the oil in a car," "I have changed spark plugs on a car," and "I have replaced brake shoes on a car." The self-rating consisted of instructions to "please place a check mark indicating your knowledge of car mechanics and repair" followed by five items and scores ranging from *very little* to *very much.*

The retention test contained the following instructions at the top of the sheet: "Please write down an explanation of how a car's braking system works. Pretend that you are writing an encyclopedia entry for people who are not already familiar with brakes." The transfer test consisted of the following four questions, each typed on a separate sheet: "What could be done be make brakes more reliable, that is, to make sure they would 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 release the pedal repeatedly and rapidly)?" The matching test presented a frame from the beginning of the animation along with the following instructions:

> Circle the brake pedal and write B next to it. Circle the piston in the master cylinder and write P next to it. Circle part of the brake line and write L next to it. Circle the smaller piston in the wheel cylinder and write W next to it. Circle part of the brake shoes and write S next to it. Circle part of the brake drum and write D next to it.

The computerized materials consisted of two computer programs for multimedia presentations on how a car's braking system works. Both programs generated an identical 45-s animation depicting a foot pressing on a brake pedal, a piston moving forward in a master cylinder, brake fluid being compressed in the brake line, pistons moving forward in wheel cylinders, pistons pressing against brake shoes, brake shoes pressing against a brake drum, and the wheel slowing down to a stop. The AN version included concurrent narration describing each of the major events in words spoken at a slow rate by a male voice. The AT version included concurrent on-screen text, using the same words and timing as the narration. The words consisted of the following 10 segments (without the segment letters): (a) When the driver steps on the car's brake pedal, (b) a piston moves forward inside the master cylinder. (c) The piston forces brake fluid out of the master cylinder, (d) and through the brake lines to the wheel cylinders. (e) In the wheel cylinders, (f) the increase in fluid pressure (g) makes a smaller set of pistons move. (h) These smaller pistons activate the brake shoes. (i) When the brake shoes press against the drum, (j) both the drum and the wheel slow down, or stop.

For Group AT, each segment appeared under the animation and stayed on until the next segment appeared; for Group AN, there was a pause after each segment. The multimedia materials were adapted from previously used paper-based passages on braking systems (Mayer, 1989; Mayer & Gallini, 1990). The multimedia presentations were created using Director 4.0 for Macintosh (Macromedia, 1994), an animation program widely available from software distributors. The apparatus consisted of five Macintosh IIci computer systems with 14-in. color monitors and Sony headphones.

Procedure. Participants were tested in groups of 1 to 5 per session. Each participant was randomly assigned to a treatment group (i.e., either AN or AT) and was seated at an individual cubicle in front of a computer.

First, participants completed the participant questionnaire, working at their own rates. Second, the experimenter presented oral instructions stating that the computer would explain how a car's braking systems works, that the computer would repeat the explanation two times, and that when the computer was finished, the experimenter would have some questions for the participant to answer. Participants in the AN treatment were told to put on headphones, and all participants were told to press the space bar to begin the presentation. Third, after pressing the space bar, the animation with narration was presented three times to participants in the AN group and the animation with text was presented three times to participants in the AT group. Fourth, when the presentation was finished, the experimenter presented oral instructions for the test, stating that there would be a series of question sheets and that for each, the participant should keep working until told to stop. Fifth, the first sheet was distributed along with instructions to write down an explanation of how a car's brakes work. After 5 min, the recall sheet was collected. Then, the problem-solving sheets were presented one at a time for 2.5 min each and each sheet was collected by the experimenter before the subsequent sheet was handed out. Finally, the matching test was presented and then collected after 2.5 min.

We used slightly different methodologies in Experiments 1 and 2 because the materials differed and because we wanted to determine whether the results would be robust. The time allocated to various tasks in each experiment was based on pilot testing.

Scoring. A mechanical experience score was computed by assigning a score of 0 to 4 on the self-assessment (with 0 points for checking *very little,* 1 point for checking the space between *very little* and *average,* 2 points for checking *average,* 3 points for checking the space between *average* and *very much,* and 4 points for checking *very much*) and adding to that 1 point for each of the six items checked on the checklist. To focus the study on participants who lacked extensive knowledge of car mechanics, data were eliminated for participants scoring above 5 on this 10-point scale and new participants were used in their places ($n = 13$).

A scorer who was not aware of the treatment condition determined the retention score, matching score, and transfer score for each participant. A retention score was computed by counting the number of major idea units (out of eight possible) that the participant produced on the retention test. One point was given for correctly stating each of the following eight idea units regardless of wording: (a) driver steps on brake pedal, (b) piston moves forward in master cylinder, (c) piston forces brake fluid out of wheel cylinders, (d) fluid pressure increases at wheel cylinders, (e) smaller pistons move, (f) small pistons activate brake shoes, (g) brake shoes press against drum, and (h) drum and wheel stop or slow down.

A matching score was computed for each participant by counting the number of correctly labeled parts (out of six possible) on the matching test. Participants received 1 point for each part that was circled and labeled with the appropriate letter.

A transfer score was computed for each participant by counting the number of acceptable answers (out of 12 possible) that the participant produced across the four transfer problems. For example, acceptable answers for the first question about redesigning brakes for reliability included a back-up system or cooling system, acceptable answers for the second question about redesigning brakes for effectiveness included using more friction sensitive brake shoes or reducing the distance between the brake shoe and pad, acceptable answers for the third question about troubleshooting a faulty brake system included the possibility of a hole in the brake line or a piston stuck in one position, and acceptable answers for the fourth question about pumping the brakes included that pumping reduces heat or reduces wearing the drum in one place. Participants received no more than 3 points per problem.

As in Experiment 1, we analyzed data from low-experience learners only. After we used the same procedure for identifying high-experience learners as in Experiment 1, 13 students were replaced in Experiment 2.

Results and Discussion

Split-attention effect on verbal recall. According to dual-processing theory, students should remember more of the verbal material when it is presented as narration than when presented as text. The retention test bars on the left side of Figure 4 show the proportion of the eight idea units correctly recalled by students in the AN group and the AT group. An ANOVA was conducted with group (AN vs. AT) as the between-subjects factor and recall score as the dependent measure. As can be seen in the retention test bars on the left side of Figure 4, there was a split-attention effect in which AN students tended to recall more relevant idea units than did AT students, $F(1, 66) = 4.630$, $MSE = 2.670$, $p < .05$. These results are consistent with the predictions of dual-processing theory.

Split-attention effect on picture-name matching. According to dual-processing theory, students should perform better on the matching test when verbal material is presented as narration than when presented as text. The matching test bars in the middle of Figure 4 show the proportion correct on the eight-item matching test by students in the AN group and the AT group. An ANOVA was conducted with group (AN vs. AT) as the between-subjects factor and matching score as the dependent measure. As can be seen in the matching test bars in the middle of Figure 4, there was a split-attention effect in which AN students tended to correctly match more items than did AT students, $F(1, 66) = 4.800$, $MSE = 2.400$, $p < .05$. These results are consistent with the predictions of dual-processing theory.

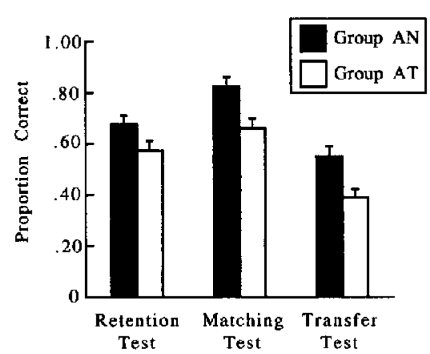

Figure 4. Proportion correct response on retention, matching, and transfer tests by two groups in Experiment 2. On retention, $M = .68$ ($SD = .19$) for Group AN and $M = .58$ ($SD = .21$) for Group AT; on matching, $M = .80$ ($SD = .26$) for Group AN and $M = .66$ ($SD = .26$) for Group AT; on transfer, $M = .55$ ($SD = .24$) for Group AN and $M = .39$ ($SD = .17$) for Group AT. AN = Group receiving concurrent auditory narration; AT = group receiving on-screen text.

Split-attention effect on problem-solving transfer. According to dual-processing theory, students should generate more problem-solving solutions when verbal material is presented as narration than when presented as text. The transfer test bars in the right side of Figure 4 show the proportion of correct solutions on the problem-solving test by students in the AN group and the AT group (with the maximum score set at 12). An ANOVA was conducted with group (AN vs. AT) as the between-subjects factor and problem-solving score as the dependent measure. As can be seen in the transfer test bars on the right side of Figure 4, there was a split-attention effect in which AN students tended to generate more solutions than did AT students, $F(1, 66) = 9.850$, $MSE = 4.520$, $p < .01$. These results are consistent with the predictions of dual-processing theory.

Overall, split-attention effects were obtained on the retention, matching, and transfer tests, yielding consistent evidence for the dual-processing hypothesis and against the information-equivalency hypothesis. On each test, the effect sizes were substantial: 0.49 for the retention test (based on $SD = 0.21$), 0.53 for the matching test (based on $SD = 0.26$), and 0.94 for the transfer test (based on $SD = 0.17$).

General Discussion

Split-Attention Effect

The major result of these studies is a split-attention effect in which students learned better when pictorial information was accompanied by verbal information presented in an auditory rather than a visual modality. The robustness of the effect was evident on three different dependent measures across two different studies.

These results extend previous research (Mousavi et al., 1995) on split-attention effects involving learning from geometry examples in a paper-and-pencil environment, in which students learned better when diagrams were accompanied by verbal explanations presented auditorily rather than visually. First, these results extend the split-attention effect from a paper-based context to an important new context, a computer-based multimedia environment. Second, the effect was obtained across three different dependent measures (i.e., retention of verbal material, matching of pictorial and verbal material, and problem-solving transfer). Third, the effect was obtained with a new type of material, namely scientific explanations of how systems work.

These results also extend previous research on contiguity effects in which students learned better when an animation depicting the workings of a scientific system and the corresponding narration were presented concurrently rather than successively (Mayer & Anderson, 1991, 1992; Mayer & Sims, 1994). In particular, the present results clarify the conditions producing the contiguity effect by showing that the advantage of presenting words and corresponding pictures at the same time depends on the modality of words.

Theoretical Implications

These results provide an important empirical test of a dual-processing theory of working memory within the domain of multimedia learning. First, according to a dual-processing theory of working memory, students learn better in multimedia environments when words and pictures are presented in separate modalities than when they are presented in the same modality. When pictures and words are both presented visually (i.e., a split-attention situation), learners are able to select fewer pieces of relevant information because visual working memory is overloaded. When words and pictures are presented in separate modalities, visual working memory can be used to hold representations of pictures and auditory working memory can be used to hold representations of words. Consistent with this analysis, across two experiments using different materials, AT students recalled fewer idea units than did AN students.

Second, in split-attention situations, the learner's attentional resources (or central executive resources) are used to hold words and pictures in visual working memory so there is not enough left over to build connections between words and pictures. In contrast, when learners can concurrently hold words in auditory working memory and pictures in visual working memory they are better able to devote attentional resources to building connections between them. Consistent with this interpretation, across two experiments, AT students made fewer correct name–picture matches than did AN students.

Third, in split-attention situations, an overload in visual working memory reduces the learner's ability to build coherent mental models that can be used to answer transfer questions. In contrast, when words are represented in an auditory working memory and pictures are represented in visual working memory, the learner is better able to organize representations in each store and integrate across stores. Consistent with this interpretation, across two experiments AT students generated fewer solutions on the transfer test than did AN students.

Practical Implications

This study calls attention to the need to broaden the goals of instructional designers. The design of multimedia presentations should be guided by the goal of presenting relevant information using words and pictures and by the goal of presenting words and pictures in a way that fosters active cognitive processing in the learner. Focusing solely on the first goal—presenting relevant information—can lead to treatments such as the one given to the AT groups in which visual working memory is likely to become overloaded. When the learner's working memory system becomes overloaded, the opportunities for active cognitive processing are reduced. Focusing on both goals—presenting relevant information in ways that promote active learning—can lead to treatments such as the one given to the AN groups in which working memory is less likely to become overloaded.

An important consideration in the design of multimedia presentations aimed at explaining how something works is whether to accompany animations with auditorily presented or visually presented words. The most important new practical implication of this study is that animations should be accompanied by auditory narration rather than by on-screen text. This implication is particularly important in light of the increasing use of animations and on-screen text both in courseware and on the World Wide Web. These results cast serious doubts on the implicit assumption that the modality of words is irrelevant when designing multimedia lessons with pictures and words.

According to a generative theory of multimedia learning (Mayer, 1997; Mayer et al., 1996; Mayer et al., 1995), active learning occurs when a learner engages three cognitive processes: selecting relevant words for verbal processing and selecting relevant images for visual processing, organizing words into a coherent verbal model and organizing images into a coherent visual model, and integrating corresponding components of the verbal and visual models. To foster the process of selecting, multimedia presentations should not contain too much extraneous information in the form of words or pictures. Thus, in the present studies, we presented highly concentrated explanations that concisely depicted the major steps in the to-be-learned process. To foster the process of organizing, multimedia presentations should represent the steps in order and with clear signals for both the verbal and visual information. To foster the process of integrating, multimedia presentations should present words and pictures concurrently in modalities that effectively use available visual and auditory working-memory resources. The major advance in this study was to identify techniques for presentation of verbal and pictorial information that minimize working-memory load—namely, accompanying visually presented animation with auditorily presented narration.

Limitations and Future Directions

The limitations of this study include that the instructional episode was short, the participants were college students, and only how-it-works material was used. Additional research is needed to evaluate the robustness of our results in more diverse settings.

These results should not be taken as a blanket rejection of the use of text captions with graphics. To the contrary, in a series of studies on text and illustrations about how devices work carried out in our lab at Santa Barbara the results consistently have shown that students learn more productively when text is presented within corresponding illustrations rather than when text and illustrations are presented on separate pages (Mayer, 1989, 1997; Mayer & Gallini, 1990; Mayer et al., 1995). Similarly, in a series of studies on worked-out geometry problem examples, Sweller and his colleagues (Sweller et al., 1990; Tarmizi & Sweller, 1988; Ward & Sweller, 1990) have shown that students learn better when text explanations are presented on the sheet with geometry problems than when presented separately. Overall, these studies provide ample evidence for the benefits of presenting short captions or text summaries within textbook illustrations.

Traditional school instruction tends to favor verbal modes of presentation, including both discourse and text. Yet, there is emerging evidence that pictorial modes of presentation, such as animation, offer a largely underused approach with high potential. To harness the power of pictorial learning, researchers need to understand how people integrate words and pictures as they engage in the act of sense-making. This study offers new information concerning how to encourage learners to integrate words and pictures in multimedia environments.

Finally, this study provides an example of how studying a seemingly applied question—the design of multimedia learning environments—can lead to advances in basic cognitive theory. Continued research on multimedia learning can contribute to an emerging theory of cognitive processing in working memory (Hitch & Logie, 1996).

References

Baddeley, A. D. (1986). *Working memory.* Oxford, England: Oxford University Press.

Baddeley, A. (1992, January 31). Working memory. *Science, 255,* 556–559.

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

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

Hitch, G. J., & Logie, R. H. (Eds.). (1996). *Working memory.* Hove, England: Erlbaum.

Macromedia. (1994). *Director 4.0* [Computer program]. San Francisco: Author.

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

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32,* 1–19.

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

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

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

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

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86,* 389–401.

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development, 43,* 31–43.

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

Paivio, A. (1986). *Mental representation: A dual coding approach.* Oxford, England: Oxford University Press.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12,* 257–285.
Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology, 81,* 457–466.
Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in structuring of technical material. *Journal of Experimental Psychology: General, 119,* 176–192.
Tarmizi, R., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80,* 424–436.
Ward, M., & Sweller, J. (1990). Structuring effective worked out examples. *Cognition and Instruction, 7,* 1–39.

Received June 20, 1997
Revision received December 1, 1997
Accepted December 1, 1997 ■

Acknowledgment of Ad Hoc Reviewers

The *Journal* wishes to thank persons who have reviewed manuscripts for this issue. (Numbers in parentheses indicate assistance with more than one manuscript.)

John W. Adams
Eric Amsel
Eric Anderman
Robert L. Bargert Drowns
Donald Biggs
Mimi Bong
Chris J. Boyatzis
Zvia Breznitz
Brent Bridgeman
Kay Bull
Donna Burns
Jim Byrnes
Marketa Caravolas
Russell Carney
Steven Ceci
Marilyn J. Chambliss
James Clark
Deborah Crown
Jeanne Day
Frank Dempster
M. Jean Dreher
Nelson R. DuBois
Jeff Epstein
Jo Ann Farver
Barbara Foorman
Lynn Fuchs
Teresa Garcia
Howard Gardner
Scott D. Gest
Art Glenberg
Priscilla Griffith
Vernon Hall
Jennifer L. Hampston
Laura Hanish (2)
Paul Hastings
John Hattie
Mary Hegarty
Kass Hogan
William G. Holliday
Kathleen Hoover-Dempsey (2)
George S. Howard
Joseph Jenkins
Tamara L. Jetton
Carla Johnson
Edward S. Johnson
William Kelemen
Melanie Killen
James Kulik
Betty Ann Levy
Marc Marschark
Dennis McInerney
William McKeachie
Michael S. Meloth
Gloria Miller
Mathew Mitchell
Peter Mosenthal
William Nagy
Richard Newman
Richard Olson
David Perkins
Thomas P. Petzel
Jan Plass
Clark Power
Gary E. Price
Kevin Quinn
K. Ann Renninger
Arthur Reynolds
Dan H. Robinson
Julia Johnson Rothenberg
J. Michael Royer
Donna Scanlon
Geoffrey Scheurman
Henri Schouwenburg
Mechthild Shaefer
Daniel Shapiro
Linda Siegel
Deborah Simmons
Richard Sparks
Hiller Spires
John Staud
Robert J. Sternberg
H. Lee Swanson (2)
Sonya Symons
Harry Triandis
Loel Tronsky
Suzanne Wade
Patrick M. Wright
Naomi Zigmond