When Is an Illustration Worth Ten Thousand Words?

Richard E. Mayer
University of California, Santa Barbara

Joan K. Gallini
Department of Educational Psychology, University of South Carolina

In three experiments, students read expository passages concerning how scientific devices work, which contained either no illustrations (control), static illustrations of the device with labels for each part (parts), static illustrations of the device with labels for each major action (steps), or dynamic illustrations showing the "off" and "on" states of the device along with labels for each part and each major action (parts-and-steps). Results indicated that the parts-and-steps (but not the other) illustrations consistently improved performance on recall of conceptual (but not nonconceptual) information and creative problem solving (but not verbatim retention), and these results were obtained mainly for the low prior-knowledge (rather than the high prior-knowledge) students. The cognitive conditions for effective illustrations in scientific text include appropriate text, tests, illustrations, and learners.

The two major media for communicating scientific information to students are words and pictures. In spite of the traditional bias toward verbal over visual forms of instruction, a growing research base suggests that text illustrations can have important effects on student learning (Levie & Lentz, 1982; Mandl & Levin, 1989; Mayer, 1989a; Willows & Houghton, 1987). This article is based on the premise that techniques for enhancing students' visual learning of scientific information represent a relatively untapped potential for improving instruction.

When is an illustration likely to improve student learning? What makes a good illustration? Which features of illustrations make them effective? How can we use illustrations to improve students' learning from expository prose? These are the kinds of questions that motivate our study. In this introduction, we examine the idea that one function of illustrations (which we call explanatory illustrations) is to help readers build runnable mental models and that several conditions must be met for explanatory illustrations to be effective.

Cognitive Functions of Illustrations

What makes a good illustration? In order to define what we mean by "good illustration" we must further specify "good for what?" Levin (1981; Levin, Anglin, & Carney, 1987) has proposed five functions of text illustrations: (1) decoration—illustrations can help the reader enjoy the textbook by making it more attractive (but without being relevant to the text); (2) representation—illustrations can help the reader visualize a particular event, person, place, or thing (such as found commonly in narrative passages); (3) transformation—illustrations can help the reader remember key information in a text; (4) organization—illustrations can help the reader organize information into a coherent structure; and (5) interpretation—illustrations can help the reader understand the text. In this article we focus exclusively on illustrations aimed at serving the interpretation function, which we call explanatory illustrations.

For this study we set our instructional goal as the promotion of learner understanding of how scientific systems work, as measured by qualitative reasoning about scientific systems (Bobrow, 1985). We define a "good illustration" as one that promotes the reader's understanding of how a scientific system works. Therefore, we focus our study on explanatory illustrations, that is, illustrations that promote interpretation processes.

Based on theories of mental models (de Kleer & Brown, 1985; Gentner & Stevens, 1983; Kieras & Boval, 1984; Larkin & Simon, 1987; White & Frederiksen, 1987), we identify two major features of illustrations that could help learners build runnable mental models: system topology and component behavior. System topology refers to the portrayal of each major component within the structure of the system. For example, the parts illustrations shown in Figures 1 and 2 present the system topologies of a brake system and a pump system, respectively. In Figure 1, the labeled components within the braking system include the tube, wheel cylinder, small pistons, brake shoe, and brake drum.

Component behavior refers to the portrayal of each major state that each component can be in and the relation between a state change in one component and state changes in other components. For example, the parts-and-steps illustrations in Figures 1 and 2 show alternative states of the components in a braking system and a pump system, respectively. In Figure 1, the parts-and-steps illustrations show the "before" and "after" states of the tube, smaller pistons, brake shoes, and wheel: the fluid in the tube can be compressed or not, small pistons can be back or forward, the brake shoe can be pressing against or away from the drum, and the wheel can be spinning or still. In addition, the relations among state changes are emphasized in the steps illustrations and the parts-and-steps illustrations: if the fluid in the tube is compressed, the smaller...
pistons move forward; if the smaller pistons move forward, the brake shoe will push against the drum; if the shoe presses against the drum, the spinning wheel will stop.

Current theories of mental models suggest the potential effectiveness of qualitative models in teaching students about scientific systems (de Kleer & Brown, 1985; Kieras & Bovair, 1984; Gentner & Gentner, 1983; White & Frederiksen, 1987). Such models are constructed to impart visual representations of how a system looks and how it functions under various state changes. A straightforward implication of this work is that experience with qualitative models will allow students to acquire alternative conceptualizations of how a system works and to develop skill in predicting the causal behavior of the system. Unfortunately, empirical research in this area is limited, but even more importantly, it generally ignores the educationally relevant issues of how illustrations can be designed and used to promote the acquisition of runnable mental models. Our line of research builds on existing theories of mental models, but focuses on the instructional question "When is an illustration most likely to be effective in promoting scientific understanding?"

Conditions for Effective Illustrations

As summarized in Figure 3, Mayer (1989b) has proposed four conditions that must be met for illustrations to be effective in promoting understanding of scientific text: (1) \textit{explanative text}—the text must present a cause-and-effect system that allows for qualitative reasoning (Bobrow, 1985); (2) \textit{sensitive tests}—the performance measures must evaluate the learner's understanding and qualitative reasoning about the system; (3) \textit{explanative illustrations}—the illustrations must help the learner build a runnable mental model of the system; and (4) \textit{inexperienced learners}—the students must not spontaneously engage in active learning processes such as the construction of a runnable mental model of the system. In summary, the text, tests, illustrations, and learners must all be appropriate for the instructional goal of fostering meaningful learning.

In this article we report on three experiments that each test six predictions generated from the model in Figure 3. In each experiment, students read a scientific text that contains explanatory illustrations, nonexplanatory illustrations, or no illustrations. Then students write down what they remember and answer questions about the text.

First, we examine two predictions concerning the test condition shown in Figure 3.

1. Explanatory illustrations improve recall of explanatory information but not nonexplanatory information. This prediction follows from schema theory (Rumelhart & Ortony, 1977): conceptual understanding depends on formulating a schema or knowledge structure about how a system operates. Specifically, the schema describes relations among parts of the system including how a change in one part of the system affects a change in another part of the system (i.e., explanatory information). Supplemental factual information (i.e., nonexplanatory information) might temporarily be stored in memory as discrete pieces of knowledge but generally would not be assimilated within the schema of the system.

2. Explanatory illustrations improve creative problem solving but not verbatim retention. This prediction is based on the idea that the most direct measure of conceptual understanding is a problem solving test that requires what Bobrow (1985) calls "qualitative reasoning about physical systems" (p. 1) or...

\footnotesize Predictions 1 through 4 are based on the idea that condition 4 is met, namely that the learners are inexperienced; therefore, data analysis concerning these predictions in each experiment was based only on data from low prior-knowledge learners.
what de Kleer & Brown (1983) call "running a causal model" (p. 158). In contrast, verbatim retention is irrelevant to the goal of meaningful understanding.

Second, we examine two predictions concerning the illustration condition shown in Figure 3.

3. Explanative illustrations are more effective in improving conceptual recall than nonexplanative illustrations. This prediction aims to isolate the features of effective illustrations, namely, only illustrations that explicitly concretize both the system topology and component behavior are likely to foster the learner's building of a runnable mental model. Learners who have built a runnable mental model are more likely to recall information about how one state change affects another but less likely to remember irrelevant facts, as compared to students who have not built mental models.

4. Explanative illustrations are more effective in improving problem-solving performance than nonexplanative illustrations. Following the same line of reasoning as presented for
Prediction 3, learners who possess runnable mental models should be able to engage in qualitative reasoning better than learners who do not (White & Frederiksen, 1987).

The final two predictions are based on the learner condition in Figure 3.

5. Explanative illustrations will improve recall of explanatory information for low prior-knowledge learners but not for high prior-knowledge learners. This prediction is based on the idea that high prior-knowledge learners come to the learning situation with a repertoire of techniques for strategically using their domain knowledge of mechanical systems (Alexander & Judy, 1988; McDaniel & Pressley, 1987), whereas low prior-knowledge learners do not. High prior-knowledge learners possess strategies for visually representing information in the text and can coordinate imagery strategies with other procedures. These techniques help high prior-knowledge learners focus on the explanatory information during learning so special illustrations are not needed. In contrast, imagery strategy research (McDaniel & Pressley, 1987) demonstrates that low prior-knowledge learners, who lack skill in visualizing on their own, will be more likely to profit from imagery-based instructional supplements than high prior-knowledge learners.

6. Explanative illustrations will improve problem-solving performance for low prior-knowledge learners but not for high prior-knowledge learners. Following Prediction 5, high prior-knowledge learners spontaneously build runnable mental models that can be used in problem solving so special illustrations are not needed.

Experiment 1 (Brakes)

Method

Subjects and design. The subjects were 96 college students recruited from the psychology subject pool at the University of California, Santa Barbara. Twenty-four students served in each of the four treatment groups: (1) the no illustrations group read a booklet about braking systems that contained no illustrations; (2) the parts illustrations group read a booklet that contained illustrations showing the major parts within each type of braking system; (3) the steps illustrations group read a booklet containing illustrations showing the major actions occurring for each type of braking system; and (4) the parts-and-steps illustrations group received a booklet with both types of illustrations. Half of the students in each group were low prior-knowledge students who rated their knowledge of automobile mechanics as “very little” and reported that they never had performed any automobile repairs, whereas the other half were high prior-knowledge students who rated their knowledge of automobile mechanics as more than “very little” and reported having performed minor car maintenance (such as changing oil or installing a spark plug).

Materials. The materials consisted of four versions of a booklet on braking systems, a subject questionnaire, and three posttests, each typed on 8.5 x 11 inch sheets of paper.

The text in the booklet was taken from the World Book Encyclopedia's (1987) entry for “brakes,” containing 750 words and 95 idea units. The text included an explanation of how each of four braking systems operate: mechanical braking (such as on bicycles, hydraulic disk braking systems (such as on the front wheels of cars), hydraulic drum braking systems (such as on the rear wheels of cars), and air braking systems (such as on trains and trucks). For each type of braking system, the text included a description of how a change in the status of one part affects changes in the status of other parts (i.e., what Mayer, 1984, has called explanatory information) as well as factual information such as historical information or descriptions of the materials used to make the parts (i.e., what Mayer, 1984, has called nonexplanatory information).

The no-illustrations version contained only the text; the parts-illustrations version added illustrations showing the status of each of the four braking systems before the brakes are applied, along with labels for the major parts (using names repeated from the text); the steps-illustrations version added illustrations showing the status of each of the four braking systems after the brakes are applied, along with a list of the major changes that occur (using wording repeated from the text); the parts-and-steps illustrations version added both the parts and the steps illustrations for each of the four braking systems. Illustrations appeared on the same page and below their corresponding paragraphs. A portion of the text is given in Table 1 and examples of each type of illustration are given in Figure 1.

The subject questionnaire solicited information concerning the students' experiences with automobile mechanics and repair. For example, students were asked to use a 5-point scale ranging from "very little" to "very much" for the item: "Please put a check mark beside the highest amount of experience you have ever had with the following types of maintenance: (a) putting air in a tire, changing a tire, changing oil, installing spark plugs, and replacing brake shoes. The recall posttest asked students to "write down all you can remember from the passage you have just read" and to "pretend that you are writing an encyclopedia for beginners."

The problem-solving posttest consisted of five questions, each printed on a separate sheet: (a) Why do brakes get hot? (b) What could be done to make brakes more reliable, that is, to make sure
Hydraulic brakes use various fluids instead of levers or cables. In automobiles, the brake fluid is in chambers called cylinders. Metal tubes connect the master cylinder with wheel cylinders located near the wheels. 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 set of smaller pistons move. These smaller pistons activate either drum brakes or disk brakes. Most automobiles have drum brakes on the rear wheels and disk brakes on the front wheels.

Drum brakes consist of a cast-iron drum and a pair of semicircular brake shoes. The drum is bolted to the center of the wheel on the inside. The drum rotates with the wheel, but the shoes do not. The shoes are lined with asbestos or some other material that can withstand heat generated by friction. When the brake shoes press against the drum, both the drum and the wheel stop or slow down.

Drum brakes consist of a cast-iron drum and a pair of semicircular brake shoes. The drum is bolted to the center of the wheel on the inside. The drum rotates with the wheel, but the shoes do not. The shoes are lined with asbestos or some other material that can withstand heat generated by friction. When the brake shoes press against the drum, both the drum and the wheel stop or slow down.

Each student's recall protocol was broken down into idea units; credit was given if an idea unit in the student's recall protocol conveyed the same meaning or contained the same key words as an idea unit in the passage (see Mayer, 1985, 1989b). The verbatim retention test was scored by tallying the number of correct answers. The problem-solving test was scored by tallying the number of correct answers produced for each problem, as described by Mayer (1989b). Examples of acceptable answers for the five problem-solving questions are: (a) Heat is caused by friction, pressure, rubbing, or pressing a stationary object (such as a brake shoe) against a rotating object (such as a rim or disk); (b) Reliability could be increased by adding a backup braking system, by using thicker shoes, tougher tubes, or stronger cables; or by developing more heat-resistant pads or shoes or a cooling system; (c) Effectiveness could be increased by injecting the brake fluid into the tubes or using faster moving fluid, by using a more friction-sensitive shoe, or by decreasing the distance between the shoe and the pad; (d) Brakes may malfunction because of lack of fluid, worn or loose pads, a jammed piston or holes in the piston, a hole in the tube or a break in the cable line, or moisture between pads and shoes; (e) Pumping the brakes reduces the build up of heat and allows the pad to press on the shoe at more than one place.

In analyzing the results, Predictions 1–6 (discussed previously) were tested. Predictions 1 and 7 examined the types of tests that are indicators of effective illustrations; Predictions

---

**Table 1**

*Portion of Text on Brakes (With Explanatory Information in Italics)*

---

**Figure 4.** Proportion correct on four posttests by treatment group for low and high prior-knowledge learners in Experiment 1.
Problem solving but not verbatim retention. The second prediction is that low prior-knowledge students who read explanatory text including explanatory illustrations (i.e., parts-and-steps illustrations) will recall more explanatory information but not more nonexplanatory information as compared to low prior-knowledge students who read text without illustrations. The top-left portion of Figure 4 shows the proportion correct response for low prior-knowledge students in each treatment group on explanatory and nonexplanatory recall. As can be seen, the parts-and-steps group outperformed the control group on recall of explanatory information (35% versus 15%) but not on recall of nonexplanatory information (19% versus 20%). Consistent with Prediction 1 and the results of Mayer (1989b), separate t-tests conducted on these data revealed that the groups differed in recall of explanatory information, \( t(22) = 4.53, p < .001 \), but not in recall of nonexplanatory information, \( t < 1 \).

Prediction 2: Explanatory illustrations improve creative problem solving but not verbatim retention. The second prediction is that low prior-knowledge students who read explanatory text including explanatory illustrations (i.e., both parts and steps) will generate more creative answers to transfer problems but will not perform better on verbatim retention of sentence wording as compared to low prior-knowledge students who read the text without illustrations. The top-right portion of Figure 4 shows the proportion correct response for low prior-knowledge students in each treatment group on problem solving and verbatim retention. As can be seen, the parts-and-steps group outperformed the control group on problem solving (43% versus 24%) but not on verbatim retention (40% versus 45%). Consistent with Prediction 2 and the results of Mayer (1989b), separate t-tests conducted on these data revealed the groups differed in problem solving performance, \( t(22) = 4.17, p < .001 \), but not in verbatim retention, \( t < 1 \).

Prediction 3: Explanatory illustrations are more effective in improving conceptual recall than nonexplanatory illustrations. The third prediction is that low prior-knowledge students who read explanatory text containing explanatory illustrations (i.e., both parts and steps) will show increases in explanatory recall over the control group, but low prior-knowledge students who read explanatory text with static illustrations (i.e., steps or parts) will not differ from the control group. The top-left portion of Figure 4 shows that the parts-and-steps group outperformed the steps, parts, and control groups on recall of explanatory information (35%, 21%, 17%, and 15%, respectively). Consistent with Prediction 3, an analysis of variance conducted on these data revealed the groups differed in recall of explanatory information, \( F(3, 44) = 9.81, MS_e = .33, p < .001 \), and supplemental Dunnett's Tests (at \( p < .05 \)) revealed that the parts-and-steps group outperformed the control group, but the other illustrations groups did not outperform the control group.

Prediction 4: Explanatory illustrations are more effective in improving problem solving performance than nonexplanatory illustrations. The fourth prediction is that low prior-knowledge students who read explanatory text containing explanatory illustrations (i.e., parts-and-steps) will show increases in problem solving over the control group, but low prior-knowledge students who read explanatory text with static illustrations (i.e., steps or parts) will not differ from the control group. The top-right portion of Figure 4 shows that the parts-and-steps group outperformed the steps, parts, and control groups on problem-solving (43%, 26%, 26%, and 24%, respectively). Consistent with Prediction 4, an analysis of variance conducted on these data revealed that the groups differed in problem-solving performance, \( F(3, 44) = 5.78, MS_e = .31, p < .002 \), and supplemental Dunnett's Tests (at \( p < .05 \)) revealed that the parts-and-steps group outperformed the control group, but the other illustrations groups did not outperform the control group.

Prediction 5: Explanatory illustrations improve recall of explanatory information for low-knowledge students but not for high-knowledge students. The fifth prediction is that the strong pattern of performance on explanatory recall obtained for low-prior knowledge students (described under Prediction 3) will be attenuated or eliminated for high-knowledge students. The bottom-left panel of Figure 4 shows the proportion correct response for high prior-knowledge students in each treatment group on explanatory recall. As can be seen, for high prior-knowledge students, the parts-and-steps, steps, parts, and control groups do not seem to differ in explanatory recall (28%, 34%, 30%, and 32%, respectively). Consistent with Prediction 5, an analysis of variance conducted on these data for high-prior-knowledge students revealed no significant differences among the groups, \( F(3, 44) = .61, MS_e = .52, ns \); in contrast, as described under Prediction 3, the explanatory illustrations were effective in improving explanatory recall for low-prior-knowledge students.

Prediction 6: Explanatory illustrations improve creative problem solving for low-knowledge students but not for high-knowledge learners. The sixth prediction is that the strong pattern of performance on problem solving obtained for low prior-knowledge students (described under Prediction 4) will be attenuated or eliminated for high-knowledge students. The bottom-right panel of Figure 4 shows the proportion correct response for high prior-knowledge students in each treatment group on problem solving. As can be seen, for high prior-knowledge students, the parts-and-steps, steps, parts, and control groups do not seem to differ in problem solving (45%, 43%, 38%, and 39%, respectively). Consistent with Prediction 6, an analysis of variance conducted on these data for high-prior-knowledge students revealed no significant differences among the groups, \( F(3, 44) = 1.17, MS_e = .22, ns \); in contrast, as described under Prediction 3, the explanatory illustrations were effective in improving problem solving for low-prior-knowledge students.

Experiment 2 (Pumps)

Experiment 2 used a passage on how pumps work and was intended to provide replicative tests of the six predictions.
Method

Subjects and design. The subjects were 96 college students recruited from the same subject pool as in Experiment 1. The design and cell sizes were identical to Experiment 1, except that high and low prior-knowledge status was based on students' experience with household repair rather than car mechanics.

Materials. Similar to Experiment 1, the materials consisted of four versions of a booklet on pumps, a subject questionnaire, and three posttests, each typed on 8.5 x 11 inch sheets of paper. The text was taken from the World Book Encyclopedia (1987) entry for "pump," containing 812 words and 128 idea units. The text included explanations of how centrifugal, sliding vane, lift, and bicycle tire pumps operate. No illustrations, parts illustrations, steps illustrations, and parts-and-steps illustrations booklets were constructed as in Experiment 1. A portion of the text is given in Table 2 and examples of each type of illustration are given in Figure 2.

The subject questionnaire was similar to that used in Experiment 1 except that it solicited information concerning the students' experience with household repair rather than automobile mechanics. For example, students were asked to use a 5-point scale ranging from "very little" to "very much" for the item: "Please place a check mark indicating your knowledge of how to fix household appliances and machines." In another item, students were asked: "Please place a check mark next to the things you have done." The list included owning a screwdriver, owning a power saw, having replaced the hoses on a lawn sprinkler system, having replaced the washer in a sink faucet, having replaced the flush mechanism in a toilet, and owning a screwdriver, owning a power saw, having replaced the washer in a sink faucet, having replaced the flush mechanism in a toilet, and owning a screwdriver, owning a power saw, having replaced the washer in a sink faucet, having replaced the flush mechanism in a toilet.

As in Experiment 1, the problem-solving posttest consisted of five questions, each on a separate sheet. The five questions were: (a) What could be done to make a pump more reliable, that is, to make sure it would not fail? (b) What could be done to make a pump more effective, that is, to move more liquid or gas more rapidly? (c) Suppose you push down and pull up the handle of a lift pump several times but no water comes out. What could have gone wrong? (d) Why does water enter a lift pump? Why does water exit from a lift pump? (e) The text you read mentioned a "screw pump that consisted of a screw rotating in a cylinder," but the text did not really explain how it works. Based on your understanding of how pumps work, please write your own idea of how you think a screw pump could be used to move water.

Table 2

<table>
<thead>
<tr>
<th>Portion of Text on Pumps (With Explanatory Information in Italics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bicycle tire pumps vary in the number and location of the valves they have and in the way air enters the cylinder. Some simple bicycle tire pumps have the inlet valve on the piston and the outlet valve at the closed end of the cylinder. A bicycle tire pump has a piston that moves up and down. Air enters the pump near the point where the connecting rod passes through the cylinder. As the rod is pulled out, air passes through the piston and fills the areas between the piston and the outlet valve. As the rod is pushed in, the inlet valve closes and the piston forces air through the outlet valve.</td>
</tr>
</tbody>
</table>

Prediction 2: Explanatory illustrations improve creative problem solving but not verbatim retention. The top-right portion of Figure 5 shows the proportion correct response for low prior-knowledge learners in each treatment group on problem solving and verbatim retention. As can be seen, the parts-and-steps group outperformed the control group on problem solving (47% versus 28%) but not on verbatim retention (66% versus 68%). Consistent with Prediction 2, the separate $t$-tests conducted on these data revealed that the groups differed in problem-solving performance, $t(28) = 3.51$, $p < .01$, but not in verbatim retention, $t(28) < 1$.

Prediction 3: Explanatory illustrations are more effective in improving conceptual recall than nonexplanatory illustrations. The top-left portion of Figure 5 shows that the parts-and-steps group outperformed the steps, parts, and control groups on recall of explanatory information (23%, 13%, 13%, and 3%, respectively). Consistent with Prediction 3 and the results of Experiment 1, an analysis of variance conducted on these data revealed the groups differed in recall of explanatory information, $F(3, 56) = 7.32$, $MS_e = .53$, $p < .001$, and supplemental Dunnett’s Tests (at $p < .05$) revealed that the parts-and-steps group outperformed the control group, but the other illustrations groups did not outperform the control group.

Prediction 4: Explanatory illustrations are more effective in improving problem solving performance than nonexplanatory illustrations. The top-right portion of Figure 5 shows that the parts-and-steps group outperformed the steps, parts, and control groups on problem solving (47%, 20%, 28%, and 28%, respectively). Consistent with Prediction 4 and the results of Experiment 1, an analysis of variance conducted on these data revealed the groups differed in problem-solving performance, $F(3, 56) = 9.20$, $MS_e = .31$, $p < .001$, and supplemental Dunnett’s Tests (at $p < .05$) revealed that the parts-and-steps group outperformed the control group, but the other illustrations groups did not outperform the control group.

Prediction 5: Explanatory illustrations improve recall of explanatory information for low-knowledge students but not for high-knowledge students. The bottom-left panel of Figure 5 shows the proportion correct response for high prior-knowledge students in each treatment group on explanatory recall. As can be seen, for high prior-knowledge students, the differences among the parts-and-steps, steps, parts, and control groups in explanatory recall (24%, 23%, 16%, and 13%, respectively) were relatively small. However, in contrast to Prediction 5 and the results of Experiment 1, an analysis of variance conducted on these data for high prior-knowledge students revealed mildly significant differences among the groups, $F(3, 56) = 2.98$, $MS_e = .53$, $p < .05$. Apparently, the emphasis on explanatory information in the illustrations influenced the recall performance of high prior-knowledge in this experiment.

Prediction 6: Explanatory illustrations improve creative problem solving for low-knowledge illustrations but not for high-knowledge learners. The bottom-right panel of Figure 5 shows the proportion correct response for high prior-knowledge students in each treatment group on problem solving. As can be seen, for high prior-knowledge students, the parts-and-steps, steps, parts, and control groups do not seem to differ in problem solving (55%, 45%, 49%, and 51%, respectively). Consistent with Prediction 6, an analysis of variance conducted on these data for high prior-knowledge students revealed no significant differences among the groups, $F(3, 56) = 5.66$, $MS_e = .72$, ns; in contrast, as described under Prediction 4, the explanatory illustrations were effective in improving problem solving for low prior-knowledge students.

Experiment 3 (Generators)

Experiment 3 used a passage on electric generators and was intended to provide replicative tests of the six predictions. This experiment differed from Experiments 1 and 2 in several important ways: the passage was longer and more complex, there were only three versions of the booklet rather than four, the parts-and-steps illustrations contained labels for parts but not for steps, the parts-and-steps illustrations booklet included...
several illustrations that were not in the same format as used in previous experiments, words were added to the parts and parts-and-steps booklets to clarify the illustrations, and prior knowledge was measured on a different kind of survey instrument.

**Method**

**Subjects and design.** The subjects were 108 college students recruited from the same subject pool as Experiments 1 and 2. The design was identical to Experiment 1 except that there was no steps group; the no-illustrations group consisted of 17 low prior-knowledge and 18 high prior-knowledge students; the parts group consisted of 15 low prior-knowledge and 21 high-prior knowledge students; and the parts-and-steps group consisted of 12 low-prior knowledge and 25 high prior-knowledge students.

**Materials** The materials consisted of three versions of a passage on electric generators, a subject questionnaire, and three posttests, each typed on 8.5 x 11 in sheets of paper.

The text for the no-illustrations booklet was taken from the *World Book Encyclopedia*’s (1987) entry on “electric generator,” containing 2066 words and 161 idea units. The parts-illustrations booklet was created by inserting 4 figures and approximately 225 words that clarified the figures. The figures showed the parts in a simplified generator, a real-life generator, a simplified AC generator, and a real-life AC generator. The parts-and-steps illustrations booklet was created by adding 13 figures and 372 words that clarified the figures. It included the same figures as in the parts booklet along with the following additions: (a) three additional illustrations of the simple generator showing the armature in each state of rotation; (b) three additional illustrations of the simplified AC generator showing the armature in each state of rotation; (c) two additional illustrations showing the relation between the simplified and real-life generator and between the simplified AC generator and real-life AC generator; and (d) a summary illustration showing the flow of current among the parts in a simplified AC generator.

The subject questionnaire consisted of 11 statements concerning the subject’s experience and interest in working with mechanical and electrical devices. Each statement was to be rated on a 5-point scale consisting of never (1), seldom (2), sometimes (3), frequently (4), and always (5). Typical questionnaire items included: “I like to repair electrical gadgets.” “I enjoy experimenting with mechanical things,” and “I like to disassemble mechanical things merely to perceive what they look like inside.”

The recall posttest was identical to that used in Experiments 1 and 2.

The problem-solving posttest contained eight questions: (a) Why is no energy generated when the armature of the generator is in a parallel position? (b) Provide reasons as to why a generator might not produce enough electrical energy. (c) What can be done to increase the energy output from an electrical generator? (d) How would you know if an AC electric generator is not functioning properly? (e) Explain why one generator might produce more cycles of current than another? (f) What happens to the electric current if no device is hooked up to the electric generator? (g) Can an AC generator ever conduct electric current in the same direction throughout one complete cycle of the rotation process? Explain your answer. (h) Suppose a radio is hooked up to an AC generator. The radio suddenly goes out. The problem seems to be related to the generator. Provide an explanation for the power failure.

The verbatim retention posttest consisted of 15 pairs of sentences such as, “An electric generator is a machine that produces electricity. An electric generator is a machine that generates electricity.

**Procedure.** The procedure was the same as in Experiments 1 and 2 except that subjects were given 25 minutes to read their booklets.

**Results and Discussion**

**Scoring.** Students’ scores on the subject questionnaire were computed by tallying the numbers (i.e., 1 through 5) circled for each of the eleven items. Students who obtained scores below 29 were counted as low prior-knowledge students and those who obtained scores of 29 or above were counted as high prior-knowledge students.

As in Experiments 1 and 2, four major dependent measures were recorded for each subject: number of explanatory idea units recalled (out of 32 possible); number of nonexplanatory idea units recalled (out of 129 possible); number of correct answers on the problem-solving test (out of 15 possible); and number of correct answers on the verbatim recognition test (out of 15 possible). The recall, problem-solving, and verbatim retention tests were scored as in Experiment 1. Examples of acceptable answers for the problem-solving questions are: (a) No energy is generated when the armature is in a parallel position because it is not cutting through the magnetic lines of force; (b) The generator might not produce enough electrical energy because the carbon brushes are not properly aligned with the slip ring, or the initial current from the prime mover is insufficient; (c) One sign of an AC generator malfunctioning is that the current does not reverse direction; (d) A generator might produce more cycles of current than another because it is larger or the armature is rotating at a faster rate; (e) If no device is hooked up to an electric generator, no electrical current is produced; (f) A generator that conducts electrical current in the same direction is not an AC generator but a DC generator; and (g) If a radio hooked up to a generator suddenly fails, the problem could be that the armature stopped rotating or the slip rings are not rotating properly.

As in Experiments 1 and 2, six predictions were tested.

**Prediction 1:** Explanatory illustrations improve recall of explanatory information but not nonexplanatory recall. The top-left portion of Figure 6 shows the proportion correct response for low prior-knowledge students in each treatment group on explanatory and nonexplanatory recall. As can be seen, the parts-and-steps group outperformed the control group on recall of explanatory information (21% versus 9%) but not on recall of nonexplanatory information (7% versus 8%). Consistent with Prediction 1, the results of Experiments 1 and 2, and the results of Mayer (1989b), separate t-tests conducted on these data revealed that the groups differed in recall of explanatory information, t(27) = 3.37, p < .01, but not in recall of nonexplanatory information, t(27) < 1, ns.

**Prediction 2:** Explanatory illustrations improve creative problem solving but not verbatim retention. The top-right portion of Figure 6 shows the proportion correct response for low prior-knowledge students in each treatment group on problem solving and verbatim retention. As can be seen, the parts-and-steps group outperformed the control group on

---

2 Several additional posttests were given but were not analyzed because they were not relevant to our six predictions.
Prediction 4: Explanative illustrations are more effective in improving problem solving performance than nonexplanative illustrations. The top-right portion of Figure 6 shows that the parts-and-steps group outperformed the parts and control groups on problem solving (56%, 39%, and 29%, respectively). Consistent with Prediction 4, an analysis of variance conducted on these data revealed that the groups differed in problem-solving performance $F(2, 40) = 21.11, p < .001$, and supplemental Dunnnett's Tests (at $p < .05$) revealed that the parts-and-steps group outperformed the control group, but the parts group did not outperform the control group.

Prediction 5: Explanative illustrations improve recall of explanatory information for low-knowledge students but not for high-knowledge students. The bottom-left panel of Figure 6 shows the proportion correct response for high prior-knowledge students in each treatment group on explanatory recall. As can be seen, for high prior-knowledge students, the differences among the parts-and-steps, parts, and control groups in explanatory recall (26%, 16%, and 11%, respectively) were relatively large. Inconsistent with Prediction 5 and the results of Experiments 1 and 2, an analysis of variance conducted on these data for high prior-knowledge students showed significant differences among the groups, $F(2, 60) = 19.96, p < .001$, and is similar to the pattern for low prior-knowledge students described under Prediction 3 in which explanatory illustrations were effective in improving explanatory recall. Thus, both in Experiments 2 and 3, this is the only inconsistency in our predictions.

Prediction 6: Explanative illustrations improve creative problem solving for low-knowledge students but not for high-knowledge learners. The bottom-right panel of Figure 6 shows the proportion correct response for high prior-knowledge students in each treatment group on problem solving. As can be seen, for high prior-knowledge students, the parts-and-steps, parts, and control groups do not seem to differ in problem solving (61%, 55%, and 55%, respectively). Consistent with Prediction 6 and the results of Experiments 1 and 2, an analysis of variance conducted on these data for high prior-knowledge students revealed no significant differences among the groups, $F(2, 60) = 1.20, ns$; in contrast, as described under Prediction 4, the explanatory illustrations were effective in improving problem solving for low prior-knowledge students.

**Conclusion**

**Conditions for Effective Illustrations**

A general goal of this study was to test the model of conditions for effective illustrations summarized in Figure 3. We examined the specific predictions that illustrations would be effective in our experiments only with explanatory text (i.e., appropriate text), tests of understanding and reasoning (i.e., appropriate tests), steps-and-parts illustrations (i.e., appropriate illustrations), and low prior-knowledge learners (i.e., appropriate learners). Table 3 summarizes whether or not each of the six predictions, based on the model of conditions for effective illustrations, was supported in each of the three experiments.
WHAT MAKES A GOOD ILLUSTRATION?

Table 3
Summary of Results: Tests of 18 Predictions

| Predictions based on four conditions for effective illustrations (in boldface) | Results of Experiment |
|---|---|---|
| **Appropriate text:** Use explanatory text (No predictions were tested.) |  |  |
| **Appropriate tests:** Use tests that measure conceptual understanding | Y | Y | Y |
| Explanatory illustrations improve explanatory recall but not non-explanatory recall. |  |  |  |
| Explanatory illustrations improve creative problem solving but not verbatim retention. | Y | Y | Y |
| **Appropriate illustrations:** Use illustrations that explain |  |  |  |
| Explanatory illustrations improve explanatory recall better than non-explanatory illustrations. |  |  |  |
| Explanatory illustrations improve problem solving better than non-explanatory illustrations. |  |  |  |
| **Appropriate learners:** Use illustrations for learners who lack prior knowledge | Y | N | N |
| Explanatory illustrations improve explanatory recall for low rather than high knowledge learners. |  |  |  |
| Explanatory illustrations improve problem-solving for low rather than high knowledge learners. | Y | Y | Y |

Note. Y = prediction was confirmed; N = prediction was not confirmed.

**Appropriate text.** The first condition in the model is that the text be appropriate for the instructional goal. Since our instructional goal was to enhance learner understanding, we used explanatory text rather than descriptive or narrative text. Although we did not directly manipulate the type of text in this study, we exclusively used text in each experiment that met the condition of appropriateness.

**Appropriate tests.** The second condition is that the text be appropriate to the instructional goal. In our experiments, we compared illustrations that concretized the changes in status of parts within the system (explanatory illustrations) to illustrations that did not (non-explanatory illustrations). The first two predictions in Table 3, relating to test appropriateness, were supported in each of the three experiments. The consistency of these results, in conjunction with similar results in two previous experiments (Mayer, 1989b), provide strong support for the idea that tests should be sensitive to the specific goals of instruction.

**Appropriate illustrations.** The third condition is that the illustrations should be appropriate for the instructional goal. In our experiments, we compared illustrations that concretized the changes in status of parts within the system (explanatory illustrations) to illustrations that did not (non-explanatory illustrations). The second two predictions in Table 3, concerning illustration appropriateness, were supported in each of the three experiments. These results provide new evidence concerning the characteristics of the models portrayed in effective explanatory illustrations: The ineffective illustrations failed to visually portray either system topology (i.e., steps illustrations) or component behavior (i.e., parts illustrations) whereas the effective illustrations (i.e., parts-and-steps illustrations) portrayed both. In particular, the parts-and-steps illustrations used a series of two (or more) frames to show the state of the components within the system at various points in the operation of the system.

**Appropriate learners.** The fourth condition is that the learners would not achieve the desired learning outcome without special instruction. In our experiments, we compared the learning outcomes of students who lacked domain-specific knowledge and those who possessed it. The final two predictions in Table 3, testing learner appropriateness, were generally supported. The major exception, that explanatory illustrations improved explanatory recall for high-knowledge learners in Experiment 3, may be accounted for by the fact that the explanatory text was expanded upon in Experiment 3 but not in the other experiments.

In summary, in the spirit of complementing Larkin & Simon's (1987) work, our goal was to answer the question, “When is a diagram worth ten thousand words?” Our results provide a four-part answer: when the text is potentially understandable, when the value of illustrations is measured in terms of learner understanding, when the illustrations explain, and when the student lacks previous experience. Finally, our report fits into a growing literature on the cognitive psychology of visual instructional methods (Holley & Dansereau, 1984; Waddill, McDaniel & Einstein, 1988; Weidenmann, 1989; Winn, 1987) and points to the potential of visually based instruction as a medium for promoting students' understanding of scientific material.

References


Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens...
(Eds.), Mental models (pp. 101–130). Hillsdale, NJ: Erlbaum.

Received September 11, 1989
Revision received May 11, 1990
Accepted May 29, 1990