Systematic Thinking Fostered by Illustrations in Scientific Text

Richard E. Mayer
University of California, Santa Barbara

In 2 experiments, students who lacked prior knowledge about car mechanics read a passage about vehicle braking systems that either contained labeled illustrations of the systems, illustrations without labels, labels without illustrations, or no labeled illustrations. Students who received passages that contained labeled illustrations of braking systems recalled more explanatory than nonexplanatory information as compared to control groups, and performed better on problem solving transfer but not on verbatim recognition as compared to control groups. Results support a model of meaningful learning in which illustrations can help readers to focus their attention on explanatory information in text and to reorganize the information into useful mental models.

What can be done to improve the understandability of expository text? In particular, how can we design text so that readers will be able to use the information creatively to solve problems? To answer these questions, my colleagues and I have been engaged in a series of studies investigating the effects of advanced organizers (Mayer, 1975a, 1976, 1978, 1979, 1980, 1983; Mayer & Bromage, 1980), signaling (Loman & Mayer, 1983; Mayer, Dyck, & Cook, 1984), and adjunct questions (Mayer, 1975b; Sagerman & Mayer, 1987). This work has been based on the idea that different instructional manipulations may have effects on different aspects of cognitive processing in different learners. In particular, assimilation theory—the idea that learning involves integrating new information with existing knowledge—suggests three primary functions of cognitive processes: to guide selective attention towards certain information in the text, to foster the building of internal connections among ideas from the text, and to foster building external connections between ideas in the text and the learner's existing knowledge (Mayer, 1982, 1984, 1985, 1987). The present study opens a new line of attack by investigating the role of illustrations as potential vehicles for helping students understand expository text.

Illustrations may help readers build useful mental model (Bayman & Mayer, 1988; Mayer, 1987). For example, Figure 1 presents an illustration corresponding to the text presented in Table 1. As can be seen, the illustration helps focus attention on the relevant information such as the state of the piston in the master cylinder and provides a context for building internal connections such as moving the piston will push fluid forward.

Evaluation of what is learned involves more than examining how much is remembered on a comprehension test. First, illustrations are useful in helping the reader to focus on or explanatory information in the text, then readers given illustrations should tend to recall more explanatory than nonexplanatory information as compared to readers not given illustrations. Second, if illustrations are successful in helping readers build connections, then readers given illustration should perform better on transfer but not on verbatim recognition as compared to readers not given illustrations. Transfer for the brakes test involves systematic thinking—building and using a mental model of the system to make predictions.

Conditions for Meaningful Learning

The research literature on text illustrations appears to be inconsistent and inconclusive (Szlachcinski, 1979; Wright, 1977, 1982). One problem with reviewing research on text illustrations is that many studies do not use appropriate research methods or consistently address theoretical issues (Levin, Anglin, & Carney, 1987). Figure 2 shows four criteria for research on meaningful instructional methods in general, and on illustrations in particular.

Potential Meaningfulness of the Material

First, the to-be-learned material must be potentially meaningful; that is, it must be possible to construct a coherent mental model from the material. If the material is not potentially meaningful, then any attempts to help students to understand it will not succeed. It follows that the selection of text materials is a crucial step in any program aimed at producing meaningful learning. Unfortunately, teachers
Table 1
Portion of Brakes Text and Examples of Internal and External Connections

<table>
<thead>
<tr>
<th>Internal Connections</th>
<th>External Connections</th>
</tr>
</thead>
<tbody>
<tr>
<td>First, you step on the brake pedal.</td>
<td>A piston in a cylinder is like a syringe.</td>
</tr>
<tr>
<td>Second, the piston moves forward in master cylinder.</td>
<td>A brake tube is like a garden hose.</td>
</tr>
<tr>
<td>Third, the brake fluid is forced through tube towards wheels.</td>
<td>Automobile disk brakes are like caliper brakes on bicycles.</td>
</tr>
<tr>
<td>Fourth, the smaller pistons move forward in wheel cylinder.</td>
<td></td>
</tr>
<tr>
<td>Fifth, the brake pads press in against disk.</td>
<td></td>
</tr>
<tr>
<td>Sixth, the rotating wheel slows or stops.</td>
<td></td>
</tr>
</tbody>
</table>

Would have to look long and hard through some textbooks to find potentially meaningful material.

To simplify matters, I have focused on explanatory text—that is, text that explains—and in particular on explanations of how systems work. A system is simply a collection of parts that interact with one another in consistent ways (Simon, 1969). In analyzing common text structures used in science textbooks, Cook and Mayer (1983, 1988) noted that systems are generally presented in the form of a sequence. In a sequence text structure, a series of state changes are listed, generally in chronological order, with the first change enabling the second change, the second change enabling the third change, and so on. The present study uses a sequence passage on vehicle braking systems, a portion of which is reproduced in Table 1.

Novice Status of the Learner

Second, the learners who have the most to gain from meaningful methods are those who would not spontaneously engage in meaningful learning processes. For example, a novice learner does not already possess or readily generate a mental model of the system. In analyzing individual differences in the effectiveness of instructional methods, Snow and Lohman (1984) have found that meaningful methods of instruction are most useful for less skilled or less knowledgeable learners. In the present study I focused on students who reported having never repaired a car.

Effectiveness of the Instructional Manipulation

Third, the instructional manipulation must effectively direct the reader's attention to the explanatory information and help the reader to build connections. The systematic illustration shown in Figure 1 is intended to accomplish these goals: It emphasizes the main elements in the system and shows how a change in the status of one element affects changes in the status of other elements.

Appropriateness of Test

Fourth, the evaluation must focus on measuring meaningful learning outcomes as evidenced by systematic thinking. In particular, systematic thinking for the brakes passage involves being able to creatively use the information in the passage to answer transfer questions, such as "How can you improve the effectiveness of car brakes?"

In summary, in the present study all four criteria for meaningful instructional method research are met: (a) A potentially meaningful passage on brakes is presented (b) to students who lack appropriate prior knowledge, and (c) the inclusion of a systematic illustration of the underlying model (d) is expected to lead to systematic thinking as evidenced by superior problem solving transfer.

![HYDRAULIC DRUM BRAKES](image)

Figure 1. Illustration of hydraulic braking system. (Adapted from World Book Encyclopedia, Volume 2 [p. 571], 1988, Chicago: World Book, Inc. Copyright 1988 by World Book, Inc. Adapted by permission.)
Experiment 1

In Experiment 1, unknowledgeable students read a text about braking systems that contained labeled illustrations or no illustrations, and then took recall, transfer, and recognition tests.

Method

Subjects and design. The subjects were 34 female college students who rated their knowledge of car mechanics and repair as "very little" and who reported that they had never performed any automobile repairs. Half of the subjects served in the illustrations group and half served in the no-illustrations group.

Materials. The materials consisted of two versions of a text on brakes, a subject questionnaire, and three posttests, each typed on 8.5-in. × 11-in. (21.6-cm × 27.9-cm) sheets of paper. The text was taken from the World Book Encyclopedia's (1987) entry for "brakes"; it contained 750 words that could be broken down into 95 idea units. The text included an explanation of how braking systems operate, including mechanical braking systems (such as on bicycles), hydraulic braking systems (such as on automobiles), and air braking systems (such as on trains and trucks). The no-illustrations version contained only the text; the illustrations versions included labeled illustrations for mechanical, hydraulic, and air braking systems. For each system, one illustration showed the key parts and another showed the changes in the status of each part when the brakes were on. Labels in the illustrations used the same wording as in the text for describing parts and actions in the systems. Illustrations appeared on the same page and below their corresponding paragraphs. A portion of the text is shown in the upper portion of Table 1 and an example illustration is shown in Figure 1.

The subject questionnaire solicited information concerning the students' experience 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 indicating your knowledge of car mechanics and repair." In another item, students were asked, "Please put a check mark next to the things you have done." The list included having a driver's license, putting air into a tire, changing a tire, changing oil, changing 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 transfer posttest consisted of five questions, each 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 they would not fail?"; (c) "What could be done to make brakes more effective, that is, to reduce the distance needed to bring a car to a stop?" (d) "Suppose you press on the brake pedal in your car but the brakes don't work. What could have gone wrong?" (e) "What happens when you pump the brakes (i.e., press the pedal and release the pedal repeatedly and rapidly)?"

The verbatim recognition posttest consisted of eight pairs of sentences; for each pair, one of the sentences had occurred verbatim in the text whereas the other was a reworded version that retained the original meaning. For example, one item on the verbatim recognition posttest was "Each car on train has its own tank of compressed air. On trains, each car has its own tank of compressed air." The instructions at the top of the sheet asked the student to "place a check mark next to the sentence that is exactly (word-for-word) identical to a sentence in the passage you read."

Procedure. Subjects were randomly assigned to treatment group. First, subjects filled out the subject questionnaire. Second, subjects were given the appropriate text to read for 5 min: Subjects in the illustrations group received both text and labeled illustrations, whereas subjects in the no-illustrations group received text without illustrations. Third, subjects were allowed 8 min to take the recall test. Fourth, subjects were allowed 10 min to take the transfer test under the experimenter's instructions subjects spent 2 min on each item and were not allowed to go back to previous items. Fifth, subjects were allowed 2 min to complete the verbatim recognition posttest.

Results and Discussion

Scoring. For scoring purposes, the text was broken down into 95 idea units using a procedure described elsewhere (Mayer, 1985). Each idea unit conveyed one main idea such as a description of a single state, event, or fact. Thirty-five of the idea units (explanatory information) described actions within the braking system that were causally related to one another, such as "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 cylinder. In the wheel cylinder, the increase in fluid pressure/ makes a set of smaller pistons move." The remaining 60 idea units (nonexplanatory information) did not systematically describe how the systems operate. For each subject a tally was made of how many explanatory (out of 35) and how many nonexplanatory idea units (out of 60) were recalled.
For each student, a transfer score was determined by tallying the number of correct answers produced on each transfer item (out of a total possible of 15). For the first question, possible elements in a correct answer included the following ideas: friction, pressure or rubbing or pressing, a stationary object (such as a brake shoe), and a rotating object (such as a rim or disk). Typical unacceptable answers lacked detail, such as saying brakes get hot because the driver rides the brakes. For the second question, some possible correct answers included adding a second or back-up system, using thicker shoes or tougher tubes or stronger cables, and developing more heat-resistant pads or shoes or a cooling system. Typical unacceptable answers involved vague statements about "the need for maintenance," such as saying that the brakes should be checked regularly. For the third question, possible correct answers included reducing the time for the brake fluid to move by injecting fluid into the tubes or using faster moving fluid, increasing the sensitivity of the shoe by using a larger shoe or extra shoe or more responsive material in the shoe, and decreasing the space between the shoe and pad. Typical unacceptable answers included pressing the pedal faster or driving more slowly. For the fourth question, some possible correct answers included lack of fluid or lack of fluid pressure, worn or loose pads or pads that do not reach the shoe, a worn or loose shoe or shoes that fall off, jammed piston or holes in piston, a hole in the fluid tube or a break in the cable line, and water or moisture between pads and shoe that reduce friction. A typical unacceptable answer was that the brakes were broken or not working properly. For the fifth question, some possible correct answers included increase in friction, reduction in the build up of heat, and pressure of pad on shoe at more than one place. A typical unacceptable answer was that pumping brakes would result in flooding the engine.

Each student's score on the recognition test was determined by counting the number of correct answers (out of a total possible of eight).

Illustrations improved recall of explanatory but not other information. The first prediction is that systematic illustrations will help direct the reader's attention towards explanatory information and away from other information in the text. The left panel of Figure 3 shows the proportion of explanatory and nonexplanatory idea units recalled by the illustrations and no-illustrations groups. As can be seen, the no-illustrations group recalled explanatory and nonexplanatory information at the same level—approximately 20%—whereas the illustrations group recalled almost twice as much of the explanatory information relative to the nonexplanatory information. An analysis of variance (ANOVA) conducted on the recall data revealed that the Group x Type interaction presented in the left panel of Figure 3 is significant, \(F(1, 32) = 4.513, MSError = .011, p < .02\). A supplemental \(t\) test conducted on the explanatory data revealed that the illustrations group recalled marginally more explanatory information than did the no-illustrations group, \(t(32) = 1.95, p < .06\).

Illustrations improved transfer but not verbatim recognition. A second prediction is that systematic illustrations would help students to build coherent mental models of braking systems. Furthermore, these models should help students answer transfer questions that require "running one's model"; however, the models should not help in verbatim retention of factual material from the text because the student may reorganize and paraphrase the material to fit with the models. The right panels of Figure 3 show the proportion of creative answers on the transfer posttest and the proportion of correct responses on the recognition posttest for each group. As predicted, the illustrations group generated significantly more creative answers to transfer questions than did the no-illustrations group, \(t(32) = 3.05, p < .01\), but the groups did not differ on recognition, \(t(32) < 1, ns\).

Experiment 2

The results of Experiment 1 encourage the idea that labeled illustrations of systems can help the reader focus attention on explanatory information and organize the information into a coherent representation of the system. A possible confound in Experiment 1, however, is that the illustrations included labels that restated the key explanatory information from the text; therefore, it is unclear whether the effects of labeled illustrations should be attributed to the labels, the graphics, or both together. Experiment 2 solved this problem by including a group that received the text with labeled illustrations added (identical to the illustrations group in Experiment 1) along with new control groups, a group that received the text with only the graphics but not the labels from the illustrations added (illustrations-without-labels group) and a group that received the text with only the labels but not the graphics from the illustrations added (labels-without-illustrations group). As in Experiment 1, all students were unknowledgeable about automobile mechanics and all took recall, transfer, and recognition tests.

Method

Subjects and design. The subjects were 44 female college students who lacked experience and knowledge in automobile mechanics. Fifteen subjects served in the illustrations group, 15 served in the illustrations-without-labels group, and 14 served in the labels-without-illustrations group.

Materials. The brakes text for the illustrations group, the subject questionnaire, recall posttest, transfer posttest, and retention posttest were identical to those used in Experiment 1. The brakes text for the illustrations-without-labels group was identical to the text for the illustrations group except that all words and pointer lines were deleted.
from the illustrations. The brakes test for the labels-without-illustrations group was identical to the text for the illustrations group except that graphics and pointer lines were deleted from the illustrations, leaving only the words used in the illustrations.

Procedure. The procedure was identical to that used in Experiment 1 except that subjects received either the illustrations text, the illustrations-without-labels text, or the labels-without-illustrations text; and subjects were given 10 min to recall the text and 2.5 min per transfer question. Scoring was identical to Experiment 1.

Results and Discussion

Labeled illustrations improved recall of explanatory but not other information. The left panel of Figure 4 shows the proportion of explanatory and nonexplanatory idea units recalled by each treatment group. If labeled illustrations serve to help the reader focus attention on explanatory information, the illustrations group should recall more explanatory than nonexplanatory information as compared with the control groups. Consistent with this prediction and the results of Experiment 1, the illustrations group recalled substantially more explanatory than nonexplanatory information whereas the two control groups did not. An ANOVA conducted on the recall data revealed a significant interaction between treatment group and type of recall that is consistent with this observation, $F(2, 41) = 5.913, MSe = .005, p < .01$. A separate ANOVA conducted on the explanatory recall data only revealed that the groups differed significantly in their recall of explanatory information, $F(2, 41) = 7.654, MSe = .008, p < .001$, and supplemental Newman-Keuls tests (using a .05 significance level) revealed that the illustrations group recalled more explanatory information than either of the control groups and that the labels-without-illustrations group recalled more explanatory information than the illustrations-without-labels group.

Labeled illustrations improved transfer but not verbatim retention. The right panel of Figure 4 shows the proportion of creative answers on the transfer posttest and the proportion of correct responses on the verbatim recognition posttest for each group. If labeled illustrations help readers to organize information into a coherent representation of the system, the illustrations group should perform better than the control groups on problem solving transfer but not on verbatim retention. Consistent with these predictions and the results of Experiment 1, the illustrations group performed more than 50% better than the control groups on creative problem solving but not on verbatim recognition. An ANOVA conducted on the problem-solving transfer data revealed that the groups differed significantly, $F(2, 41) = 14.097, MSe = .019, p < .001$, and supplemental Newman-Keuls tests revealed that the illustrations group performed significantly (at $p < .05$) better than each of the other groups. An ANOVA conducted on the verbatim retention data revealed no significant differences among the groups, $F(2, 41) < 1, MSe = .011, ns$.

Conclusion

The results are consistent with the idea that illustrations can affect the cognitive processing of the reader. In particular, the labeled illustrations used in these studies seem to have helped our students guide selective attention and build internal connections. The effects of labeled illustrations on guiding attention is indicated by the illustrations group recalling more explanatory than nonexplanatory information relative to the control groups in both Experiments 1 and 2. The effects of labeled illustrations on building connections is indicated by the illustration group outperforming the control groups on problem-solving transfer but not verbatim retention in both Experiments 1 and 2.

The results of Experiment 2 help to clarify that the effects of labeled illustrations depend both on the graphics and the labels. Providing only pictures (without corresponding labels) or only labels (without corresponding pictures) did not allow students to build useful mental models of the system as indicated by problem solving transfer, whereas students given labeled illustrations performed much better. Not surprisingly, providing labels that repeated explanatory information from the text helped students recall that information better than if no labels were presented; however, without a coherent diagram that integrated the information, students performed relatively poorly on problem-solving transfer. These results help to clarify further Larkin and Simon’s (1987, p. 65) analysis of “why a diagram is (sometimes) worth a thousand words.”

The present results instantiate the general criteria for successful instructional methods for meaningful learning described in the introduction. The criteria involve the type of text, the type of learner, the type of instructional method, and the type of performance evaluation.

Explanative Text

The first prerequisite for meaningful learning is that the to-be-learned material is potentially meaningful. Expository text can vary from purely descriptive passages whose purpose is to present a series of facts to explanatory passages whose purpose is to present a model of some system. The present study used an explanatory passage, that is, a passage that provided explanations. In particular, the passage explained how braking systems work, including mechanical, disk, drum, and air brakes. The explanation involved a prose structure that we have called “sequence” (Cook & Mayer, 1988), in which i

![Figure 4](image-url)  
Figure 4. Proportion correct on recall, transfer, and recognition tests for three groups: Experiment 2.
change in one part of a system causes a change in a connected part, which causes a change in the next connected part, and so on. In summary, the brakes passage used in Experiments 1 and 2 met the first criterion of potential meaningfulness.

Less Knowledgeable Learners

The second prerequisite is that the students would not normally engage in meaningful learning. Students can vary from those with strong learning skills and high prior knowledge to those with low prior knowledge and weak learning skills. The present study involved students who reported being unknowable about automobile mechanics and who had never performed any automobile repairs such as changing oil, spark plugs, or brake shoes. The rationale for focusing on less knowledgeable learners is that they are less likely than high-knowledge students to spontaneously engage in meaningful learning when presented with a passage involving automobile mechanics. In summary, the students in Experiments 1 and 2 met the second criterion of low knowledgeability.

Systematic Illustrations

The third prerequisite is that the instructional manipulation—in this case, illustrations used in the text—evokes meaningful learning. Illustrations in text can range from irrelevant photographs to systematic illustrations of the relations among changes in system parts. The present study used systematic illustrations of the status of braking system before and after the brake pedal is engaged, showing the relation between a change in one part of the system and a corresponding change in the next part. The rationale for focusing on systematic illustrations is that they help readers focus attention on the explanatory information and to organize that information into a coherent representation of the system. In summary, the instructional manipulation met the third criterion of evoking meaningful learning.

Transfer Appropriate Testing

Finally, the fourth prerequisite is that the performance tests measure meaningful learning. Tests can range from retention of facts to creative problem solving transfer. The present study used tests for remembering explanatory information (to evaluate whether students focused their attention) and for creatively solving transfer problems (to evaluate whether students built logical connections). In summary, the testing procedure met the fourth criterion of evaluating meaningful learning.

These results suggest some changes in the kinds of research questions we should ask about meaningful methods of instruction. Instead of asking "Do illustrations improve learning?" a more productive question is "Which kinds of illustrations affect which cognitive processes as measured by which dependent measures for which students reading which types of text?" The present results suggest that systematic illustrations can help unknowerable readers to focus attention on explanatory information in text and build connections. Finally, these results extend some of the classic work on the role of pictures in learning of low-level paired-associates (Lumsdaine, 1963; Wittrock, 1986) to the learning of higher level material.

References


Received July 25, 1988
Revision received October 20, 1988
Accepted November 15, 1988