Learning From Text With Diagrams: Promoting Mental Model Development and Inference Generation

Kirsten R. Butcher
University of Colorado at Boulder

Two experiments investigated learning outcomes and comprehension processes when students learned about the heart and circulatory system using (a) text only, (b) text with simplified diagrams designed to highlight important structural relations, or (c) text with more detailed diagrams reflecting a more accurate representation. Experiment 1 found that both types of diagrams supported mental model development, but simplified diagrams best supported factual learning. Experiment 2 replicated learning effects from Experiment 1 and tested the influence of diagrams on novices’ comprehension processes. Protocol analyses indicated that both types of diagrams supported inference generation and reduced comprehension errors, but simplified diagrams most strongly supported information integration during learning. Visual representations appear to be most effective when they are designed to support the cognitive processes necessary for deep comprehension.

Keywords: learning, diagrams, self-explanation, mental models, comprehension processes

As multimedia technology becomes increasingly popular in formal and informal educational settings, the importance of research investigating learning with visual and verbal materials takes on added value. Understanding the ways in which visual materials influence learning will be essential to developing multimedia tools with consistent and predictable benefits. Advancing technology has meant that multimedia often now includes complex forms of interactive and computationally intensive presentations; however, multimedia can be more simply defined as any presentation that includes verbal and visual information (Mayer, 2001).

In practice, basic types of multimedia—such as pictures and text—still appear to be frequently used. Currently, many digital and print materials use pictures, diagrams, and text as their primary communication format. But how does the visual representation of information influence learning? Can changes in comprehension processes account for the impact of diagrams on learning? The purpose of this research was to investigate potential effects of different diagram representations on students’ learning outcomes and comprehension processes when diagrams were added to a science text.

Learning With Text and Diagrams

Early research on pictures and text consistently demonstrated that students learned more after reading illustrated versus non-illustrated text (for a review, see Levie & Leniz, 1982). These studies predominantly administered memory measures for the source materials, including multiple-choice and fill-in-the-blank tests. But a long history of cognitive research has distinguished between rote memorization and deeper understanding (e.g., Bransford & Franks, 1971; Hilgard, Irvine, & Whipple, 1953; Olander, 1941; Tyler, 1959; see Kintsch, 1998, for a discussion). Deeper learning—evidenced by measures that assess application and transfer of information—was not widely tested in early research on illustrations. However, one early set of studies (see Dwyer, 1967, 1968, 1975) did test both memory and deep comprehension for an illustrated text. Dwyer (1967, 1968, 1975) found benefits for illustrated text when testing visual production (having participants draw a diagram and label diagram components after learning), component identification, and multiple choice but found no comprehension advantage when comparing the illustrated text with text-only conditions.

The lack of a comprehension advantage with diagrams in Dwyer’s (1967, 1968, 1975) early studies can be contrasted with the more recent body of research on multimedia comprehension developed by Mayer and his colleagues (for a summary, see Mayer, 2001). In a large number of studies investigating multimedia effects on learning, this research has found a general (although not completely unanimous) advantage for memory (e.g., Mayer, 1989; Mayer & Gallini, 1990) and an overwhelming and consistent advantage for deep comprehension (e.g., Mayer & Anderson, 1992; Mayer & Gallini, 1990) when students receive a multimedia (verbal and visual) presentation as opposed to a text-only presentation. Mayer (2001, 2003) referred to these benefits as the multimedia effect. In these studies, memory effects are indicated by performance on recall or recognition questions, and comprehension performance is measured by transfer questions that draw on the learning material but whose answers are not explicitly ad-
dressed; transfer questions require students to generate new solutions or applications based on their learning.

Consistent comprehension benefits found in recent multimedia research highlight the progress that has been made in identifying effective forms of visual and verbal information for learning. A variety of multimedia learning studies by Mayer and his colleagues (discussed in Mayer, 2001) have demonstrated the importance of carefully designed multimedia presentations and have generated a collection of important principles that predict useful multimedia conditions for learning. These principles include spatial contiguity, temporal contiguity, coherence, modality, redundancy, and individual differences (see Mayer, 2001). A thorough discussion of these principles is beyond the scope of this work, but Mayer and his colleagues (e.g., Mayer, 2001) have repeatedly found improved learning performance when multimedia principles are followed rather than violated. As discussed later, the multimedia principle most relevant to this work is the coherence principle. This principle states that learning materials should include only relevant multimedia and should avoid irrelevant pictures, sounds, or words (Mayer, 2001; Mayer, Heiser, & Lonn, 2001; Moreno & Mayer, 2000).

Multimedia principles can be used to optimize learning by maximizing the effectiveness of multimedia design. These principles help govern the choices of multimedia content—for example, in choosing relevant rather than irrelevant pictures—as well as its presentation format. However, principles governing the format of individual elements in learning materials have not been as strongly or consistently identified. For example, what type of diagram representation best supports learning from text?

What Makes Diagrams Difficult to Understand?

Previous research has suggested that different visual representations are not equally effective for all learners. Hegarty, Carpenter, and Just (1991) proposed that diagrams become more difficult to interpret as they become increasingly schematic. Schematic diagrams often depict abstract relationships (as in the case of Venn diagrams or flowcharts) and, thus, do not preserve the physical relationships present in the source information. Schematic diagrams can be difficult to interpret because the reader or learner must be able to understand and make use of abstract visual conventions to interpret them correctly. Indeed, previous research has demonstrated the importance of experience in using abstract representations effectively. For example, Petre and Green (1993) studied the use of electronic circuit schematics—diagrams that make heavy use of abstract conventions such as symmetry and proximity for functional association—by novice and expert electronics designers. Petre and Green’s results overwhelmingly indicated that, unlike experts, novices were unable to make use of abstract notation such as logical groupings and also were unable to determine what was important versus irrelevant in the schematic representations.

Compared with schematics, iconic diagrams are less abstract and usually depict a close correspondence between the diagram and the concrete object that it is intended to represent (Hegarty et al., 1991). In iconic diagrams, structural and relational information between components is central, and, as a result, iconic diagrams rely less on knowledge conventions for their interpretation. However, the structural relations in iconic diagrams are not necessarily transparent to all learners; the correspondence between an iconic diagram and the physical object it depicts means that increasing complexity in a physical object requires increasing complexity in an iconic diagram. Simplification of this complexity may result in additional abstraction that removes some of the correspondence between the diagram and its physical object, but it is likely that additional abstraction will benefit the learner when simplification makes relevant components of the diagram more clear or easily identified.

Diagram Complexity

Although many learners might assume that a faithful depiction of an object would be most useful for learning, previous research has demonstrated that the use of realism is not always a benefit in iconic diagrams. Parkhurst and Dwyer (1983) found that additional realism (as depicted in shaded drawings of the human heart) hindered low- and medium-IQ participants’ learning outcomes. This effect may be due to the possibility that lower ability learners have difficulty processing richer and possibly redundant information (Winn, 1987). However, an overall conclusion is complicated by other studies showing that less able students perform better when some elements of pictorial realism are added to diagrams (Holliday, Brunner, & Donais, 1977; Moyer, Sowder, Threadgill-Sowder, & Moyer, 1984).

Discrepancies in research results likely stem from difficulty in determining the amount of appropriate complexity and realism in diagrams for lower knowledge learners. To some extent, complexity and realism are confounded in diagrams because adding realistic detail almost always increases diagram complexity. Conversely, simplifying iconic diagrams often involves introducing some form of abstraction—in structure or function—to the concrete object being represented. As has been suggested by other researchers (e.g., Kalyuga & Anderson, 1990; Larkin & Simon, 1987; Seel & Strittmatter, 1989), the power of iconic representations may be that they make structural relations explicit, but learners with varied knowledge may not extract structural relations from complex iconic diagrams equally effectively. Thus, increased detail may help lower knowledge learners when they have difficulty extracting relevant diagram information, but the same detail may hinder them when it is not helpful in locating relevant visual information or in matching diagram elements to text.

Because comprehension of text with diagrams requires the selection of relevant components, the integration of related verbal and visual information, and the representation of such an integration (e.g., Kalyuga, Chandler, & Sweller, 2000; Mayer, 2001), it is likely that the representational complexity of diagrams plays an important role during comprehension. Thus, it was hypothesized that making a diagram’s structural relations more explicit (by abstracting structural information in the service of a functional explanation) would support learning in lower knowledge students better than a fully explicit diagram.

Reduction of a diagram’s representational complexity can be seen as an extension of the multimedia principle of coherence proposed by Mayer and his colleagues (Mayer et al., 2001; Moreno & Mayer, 2000). The coherence effect summarizes research findings showing that student learning is supported when extraneous materials are removed from multimedia lessons. Extraneous materials include interesting but informationally irrelevant words, pictures, sounds, and music. For example, video clips of a lightning storm embedded in a lesson on how lightning forms hinders
detailed diagrams more faithfully reflect the heart’s anatomical
omy to highlight functional workings of the heart, whereas the
diagram emphasizes. Simplified diagrams reflect modified anat-
diagrams, reflecting the type of information that each
structural
appropriate terms for the diagrams could be
definition of representational complexity. An alternative set of
relative difference between the diagrams rather than an absolute
diagrams, but it should be noted that this reflects a choice in
system (see Figure 1).

In the current experiments, I refer to simplified and detailed
diagrams, but it should be noted that this reflects a choice in
terminology. The reader is cautioned that these labels reflect the
relative difference between the diagrams rather than an absolute
definition of representational complexity. An alternative set of
appropriate terms for the diagrams could be functional versus
structural diagrams, reflecting the type of information that each
diagram emphasizes. Simplified diagrams reflect modified anat-
omy to highlight functional workings of the heart, whereas the
detailed diagrams more faithfully reflect the heart’s anatomical
structure. The diagrams could also be described by terms based on
research using the STEAMER training system (Hollan, Hutchins,
& Weitzman, 1984), where conceptual fidelity reflects the extent
to which a simulation depicts the mental models that are needed
to understand and to use a system, and physical fidelity emphasizes
its actual, physical characteristics. In the current experiments,
simplified diagrams are more tied to conceptual fidelity, whereas
detailed diagrams reflect more physical fidelity.

Diagram Representation and Learner Background
Knowledge

Although simplifying the representation of structural relations in
a diagram should simplify its interpretation, it is not always the
case that easier or simpler materials always benefit learners. The
body of research concerning background knowledge and learning
materials suggests that a diagram’s influence on learning may
depend on both the complexity of the diagram and the prior
knowledge of the learner. Kalyuga et al. (2000) found that the form of multimedia repre-
sentation that is most useful for learners can change as participants
become more knowledgeable in a domain. In fact, Kalyuga, Ayres,
Chandler, and Sweller (2003) found that expert students benefited
most from the same presentation format that was most unhelpful to
novice students (and vice versa); highly trained students benefited
most from a visual-only presentation and least from a presentation
with ample text, whereas novices showed the opposite pattern. Kalyuga et al. (2003) referred to the changing pattern of optimal
materials based on a learner’s knowledge as the expertise reversal
effect.

Figure 1. a: A simplified diagram, depicting blood flow from the atria
into the ventricles. b: A detailed diagram, depicting the same process.

The finding that the optimal format of visual material may
depend on a learner’s background knowledge is consistent with
findings from text comprehension research, where a learner’s prior
knowledge has been found to have clear impact on learning from
different text materials. Text comprehension research has shown
that readers with adequate domain knowledge benefit from more
difficult texts; whereas novices need clear and explicit materials,
more expert students often need additional challenges to fully
engage in the learning task. For example, unlike low-knowledge
learners, high-knowledge readers learn more (as indicated by sit-
uation model measures) from low-coherence texts (McNamara,
Kintsch, Songer, & Kintsch, 1996; McNamara & Kintsch, 1996).
The body of previous research on background knowledge and the optimal format of learning materials suggests that the usefulness of certain diagrams may depend on a learner’s background knowledge, with higher knowledge learners requiring more difficult materials in order to optimize their learning, and lower knowledge learners requiring materials that are more explicit and clear. Thus, it was hypothesized that high-knowledge participants would benefit most from the detailed diagrams and that low-knowledge participants would benefit most from the simplified diagrams. In the current research, student background knowledge is measured by the accuracy of students’ existing mental models of the heart and circulatory system. Experiment 1 investigates the effects of diagram representation and learner background knowledge on learning outcomes, including mental model development, memory for information, and transfer of knowledge.

Whereas Experiment 1 seeks to explore aspects of the multimedia effect, Experiment 2 delves deeper into cognitive reasons underlying the multimedia effect by assessing potential changes in comprehension processes resulting from the use of diagrams during text learning. Specifically, Experiment 2 was designed to determine how diagrams might affect the processes that students perform during learning. Currently, little is known about the ways in which diagrams affect comprehension processes and whether a diagram’s representation may influence these processes. In order to directly measure the influence of diagrams on cognitive processes, Experiment 2 uses self-explaining methodology (e.g., Chi, 2000; Chi & VanLehn, 1991) and verbal protocol analyses to assess the influence of diagrams on student learning.

**Self-Explaining**

Self-explaining has been used extensively by Chi and her colleagues (for a summary, see Chi, 1997) and refers to a type of verbal protocol in which participants explain the content of a text during learning. The goal of the explanations is for students actively to make sense of what they are learning (Chi, 2000). To clarify terminology, the term self-explaining is used to refer to the process of generating statements during learning by explaining the content of learning materials to oneself, self-explanation is used to refer to a single statement generated when self-explaining, and self-explanation inferences refer to knowledge inferences made when students are self-explaining (Chi, 2000). The content of verbal protocols generated during self-explaining represents a learner’s knowledge in the context of that learner’s mental representation of the problem or domain. For example, by analyzing the content of self-explanations generated during learning, Chi and VanLehn (1991) were able to analyze what knowledge gave students the ability to solve physics problems and to understand the knowledge from which students generated self-explanations. Thus, self-explanations can create a picture of what knowledge is important in comprehension and can indicate how and when inferences are derived.

It is the characterization of prior knowledge, the use of such knowledge during learning, and the ability to represent comprehension processes during learning that makes self-explaining an ideal method for testing the cognitive impact of diagrams during text comprehension. Analysis of verbal protocols generated during self-explaining can provide a window on the comprehension processes in which students engage while learning. Tracking the generation of self-explanation inferences, in particular, offers a unique method for determining whether diagrams promote inference generation during learning and for investigating the potential influence of diagram representation on these processes.

**Experiment 1**

The primary goal of Experiment 1 was to investigate whether diagram representation would interact with an individual’s background knowledge to predict learning outcomes. A second goal of Experiment 1 was to establish a multimedia effect using the current materials consistent with results from prior research (e.g., Mayer, 2001, 2003).

Experiment 1 consisted of three stages: (a) assessment of participants’ existing knowledge using two techniques; students drew and explained what they knew about the heart and circulatory system and then completed a written pretest assessing general (factual) knowledge of the domain. (b) Participants learned about the domain using one of three possible types of online materials presented as a series of Web pages: (1) text only, (2) text with simplified diagrams, or (3) text with more detailed diagrams. (c) Assessment of participants’ learning outcomes in four areas: Students drew and explained what they knew about the heart and circulatory system, completed a written posttest of general knowledge (same as pretest), answered memory questions about the text, and completed inferences questions related to the text.

**Method**

**Participants**

Participants were 74 undergraduates from the University of Colorado at Boulder. All were native English speakers and received partial credit in an introductory psychology class for their participation.

**Learning Materials**

All participants read the simplest text about the heart and circulatory system used by Wolfe et al. (1998); this text (excerpted from Silverstein & Silverstein, 1983) was written at an elementary level and consisted of 1,616 words. For this research, the text was broken into 43 HTML pages presented in Internet Explorer, typically with a 1–4 sentences on each page. Learners controlled the pace of the presentation by clicking an arrow at the bottom of each screen to continue to the next page. In the diagram conditions, 32 pages included a relevant diagram adjacent to the text. Each diagram depicted the text information that it accompanied; diagrams were modeled on the “How Your Heart Works (Lower Elementary): The Heart” handout distributed by the American Heart Association (1988) Schoolsite Program. A series of simplified and detailed (see Figure 1) iconic diagrams were produced that depicted concepts from the experimental text. The detailed diagrams preserved the anatomical structure of the original diagram, but the simplified diagrams modified the heart’s anatomy in order to more clearly depict the functional workings of the heart and the relationships between heart components.

**Learning Outcome Measures**

**Mental model improvement: Proportion of possible gain.** This measure represents the degree to which participants in each experimental condition were able to improve their existing mental models of the heart and circulatory system. Five mental model categories (explained below) were used in this experiment, and each was assigned a score from 0 to 5 corresponding to its complexity and accuracy: 0 = the least accurate no loop model and 5 = the complete and accurate double loop 2 model. Proportion of possible mental model improvement was calculated as the
difference between the student’s mental model scores pre- and postlearning divided by the maximum possible increase in mental model score.

**Categorization of mental models.** Participants were asked to draw a picture of what they knew about how the heart and circulatory system works, explaining as they drew. Drawings were categorized according to the mental model (Chi, 1997; Chi, de Leeuw, Chiu, & La Vancher, 1994) of the heart and circulatory system that they depicted. Drawings were completed before and after learning; premodels refer to students’ mental models diagnosed before learning, and postmodels refer to students’ mental models diagnosed after learning. Two raters compared students’ mental model drawings and verbal explanations with the list of necessary features for each mental model developed by Chi et al. (1994). Initial rater agreement calculated by weighted Kappa ($k_{\text{premodels}} = .97$, $k_{\text{postmodels}} = .85$) was very good, but disagreements were resolved through discussion until 100% agreement was reached.

The models identified by Chi et al. (1994), from least to most advanced, are as follows: (a) no loop, (b) ebb and flow, (c) single loop, (d) single loop with lungs, (e) double loop 1, and (f) double loop 2. Mental models are briefly described here (see Chi et al., 1994, for a detailed discussion) and are illustrated by student drawings (see Figure 2).

The critical feature of the no loop model is that blood flows from the heart to the body but does not return to the heart. The participant who drew the no loop model in Figure 2 described blood as “gushing” out of the heart to the body and decided that used blood does not return to the heart.

The ebb and flow model (see Figure 2) reflects knowledge of blood vessels as well as the fact that blood returns to the heart from the body. However, the crucial feature of the ebb and flow model is that blood travels to and from the heart via the same blood vessel(s).

The single loop and the single loop with lungs (see Figure 2) models share a key feature: Blood makes one and only one continuous loop with the heart and the body. For the single loop with lungs model, the continuous loop includes a path through the lungs to oxygenate the blood before or after traveling through the rest of the body.

The double loop 1 model (see Figure 2) correctly reflects two loops of blood flowing from the heart—one loop flowing to and from the body and another loop flowing to and from the lungs—but the path of blood flow in this model is erroneous. For example, the double loop 1 model in Figure 2 depicts blood returning from the body to the bottom left of the heart (instead of the top right) and blood flowing across the heart in two directions (instead of from top to bottom).

Finally, the double loop 2 model (see Figure 2) reflects a complete and accurate understanding of the path of blood through the heart and circulatory system. Medical conventions dictate that left and right in diagrams are reversed from the viewer’s perspective (as if looking at a person facing the viewer). The double loop 2 drawing in Figure 2 depicts the accurate path of blood: Blood exits the left ventricle (lower chamber), flows to the body, and then returns to the right atrium (upper chamber). From the right atrium, blood flows to the right ventricle and then to the lungs where carbon dioxide in the blood is exchanged for oxygen. Oxygen-rich blood is returned to the left atrium, flows to the left ventricle, and, again, flows to the body through the aorta. Blood always flows from top to bottom in the heart and cannot cross sides within the heart. Each of these elements was necessary for a participant’s drawing and explanation to be categorized as a double loop 2 mental model.

Two students produced a no loop premodel, and 2 students produced an ebb and flow premodel; these students were not included in analyses because their premodels were so infrequent. In addition, 3 participants produced a double loop 2 premodel and were excluded because they could not improve their mental models. Thus, data from 67 of the 74 participants were analyzed.

Analysis of verbal protocols prompted the collapse of the single loop mental model categories in the participant sample. A number of participants with single loop premodels knew that the lungs oxygenated blood but believed that the lungs sent oxygen into the heart. Separating participants with functionally equivalent knowledge about oxygen into separate categories did not appropriately describe their overall mental model for this experiment. Indeed, the pattern of results for single loop and single loop with lungs participants did not differ, and, for simplicity, these participants were grouped into a single category called single loop premodels.

**Pre- and posttest of general (factual) knowledge about the domain.** The General Knowledge Test from Wolfe et al. (1998) was used to assess each participant’s factual knowledge of general information about the human heart and circulatory system. As in Wolfe et al., this measure was scored as a proportion of possible gain from the pre- to the posttest, but only questions with a visual component (e.g., structure, location) were analyzed in this experiment, for example, “How many chambers are there in the heart?” Because one question (“How many continuous, closed circuits of blood are there from the heart?”) could be answered for a total of 3 points by giving the correct number of circuits and naming them (two circuits: pulmonary–lungs and systemic–body), the pre- and posttest of

---

Figure 2. Mental model categories, as drawn by participants.
general knowledge consisted of eight questions answerable for a maximum score of 10 points. Additional questions are provided in Appendix A.

Memory questions about the text. These questions were administered only after learning and addressed specific details from the text about the heart and circulatory system. Questions were selected to include details that most participants would be unlikely to know or to be able to guess without having read the text. For example, “How many times will the average heart beat during a lifetime?” Additional examples are shown in Appendix A. Some questions (e.g., “Where are the valves in the heart located?”) asked participants to provide multiple pieces of information in their answers. Thus, a total of eight text-specific memory questions were answerable for a maximum score of 17 points.

Inference questions related to text. These questions also were administered after learning and required the participant to integrate information found in the text and to apply such information to new situations or problems; answers were not addressed explicitly in the learning materials. For example one question asked, “What would be the consequences of a large hole in the septum that separates the left and right ventricles?” Correctly answering this question would require the participant to recognize that the septum exists to separate the oxygenated and deoxygenated blood (a fact not explicitly addressed in the text) and to apply this structural information to the learned concepts of energy production and carbon dioxide removal. Full integration of the inference and the learned information would allow the participant to reason about how a hole in the septum would affect essential processes (e.g., less oxygen in the blood means less energy, too much carbon dioxide can kill body cells) and to explain the resulting effects on the human body (fatigue, possible death, etc.). Additional inference questions can be found in Appendix A. Again, multipart answers meant that a total of five inference questions were answerable for a maximum score of 11 points.

Background knowledge categories. Because single loop premodels reflect a low-level understanding of the heart and circulatory system, participants with these premodels were defined as having low background knowledge. In contrast, students with double loop 1 premodels have a more advanced understanding of the heart and circulatory system but still can improve their mental models; these participants were considered to have high background knowledge. Prior knowledge (low vs. high) was used as an independent variable in Experiment 1 analyses to investigate the influence of prior knowledge on diagram effectiveness and learning outcomes.

Procedure

Assessing existing knowledge. A drawing and explaining task was chosen to assess students’ existing understanding of the heart and circulatory system because the number and type of facts participants know about a domain does not necessarily represent their integrated mental model of the domain (Chi et al., 1994). Participants were tested individually and were given blank paper and a set of pens and markers. They were asked to draw and to explain aloud everything they knew about the heart, its anatomy, and how blood moved through the heart and the rest of the system. All explanations were recorded for later transcription. If necessary, the experimenter asked for clarification when drawings or explanations were unclear, but there were no differences in the number of experimenter prompts given to students by experimental group or by participants’ existing premodels (Fs < 1). After drawing their premodels, participants completed a computerized (short-answer) pretest of general knowledge about the heart and circulatory system. Common gateway interface scripts were used to generate HTML forms that presented each question with a textbox for student answers. Participant data were automatically saved on submission as the next set of questions was returned by the server; all participants were given the same questions in the same order. Participants were asked to include any and all information that could be relevant to the questions and to guess if they were unsure of an answer. The total time allotted for assessing existing knowledge was 20 min, but student completion of activities was self-paced.

Learning phase. Participants were randomly assigned into one of the three experimental conditions: text only, simplified diagrams, or detailed diagrams. They were instructed to study the materials carefully and were informed that they would not be able to go back to previous Web pages during learning. Progress through the materials was self-paced; exact study time was not recorded, but students generally were able to complete the learning phase in about 20 min.

Postknowledge assessment. To begin the assessment phase, participants again drew and explained what they knew about the heart and circulatory system. This posttest drawing procedure was identical to the pretest drawing task, except that it occurred after learning. Analyses of the experimenter’s prompts for clarification during drawing are presented in the Results and Discussion section.

After completion of their postmodels, participants were asked to answer a series of questions. Again, participants were asked to include all relevant information and to guess if they were unsure of an answer. First, students saw the posttest of general knowledge presented on the computer; this set of questions was identical to the pretest of general knowledge. Participants had been informed that they would answer questions after learning, but they had not been told that they would retake the pretest in addition to answering new questions. When participants completed the general knowledge questions, the computer automatically moved them to the new (posttest only) questions: first to the memory questions and then to the inference questions. Forms seen by participants did not indicate the change in question type, and all participants saw the same questions in the same order. Progress through all assessments was self-paced, but students generally completed the questions within 40 min.

Results and Discussion

Analysis of variance (ANOVA) comparisons were conducted using single degree-of-freedom contrasts in Proc GLM (SAS Institute, 1990) to assess a priori hypotheses, specifically: (a) a knowledge effect such that high-knowledge learners would outperform low-knowledge learners, (b) a multimedia effect such that learners who used diagrams would outperform learners using only text, (c) a simplified diagram (quadratic) effect such that simplified diagrams would be more effective than text or detailed diagrams (this hypothesis tests whether data are accurately represented by a quadratic curve where participants using simplified diagrams outperform those using detailed diagrams and text only), and (d) a knowledge and diagram interaction such that background knowledge (high vs. low) would interact with diagram complexity (simplified vs. detailed). This hypothesis tests whether the shape of the quadratic curve differs for high- and low-knowledge students. Table 1 presents the means and standard deviations for analyses of Experiment 1 learning outcomes. Experiment 1 learning outcomes.

Table 1 presents the means and standard deviations for analyses of Experiment 1 learning outcomes. Experiment 1 results appear in the left-hand column. As seen in Figure 3, most assessments appear to show an advantage for participants using the simplified diagrams, especially for lower knowledge learners.

Mental model improvement. Results yielded a main effect of knowledge level, $F(1, 66) = 4.9, p = .031, \eta^2_p = .07$; high-knowledge participants ($M = .88, SD = .34$) improved their
mental models more than lower knowledge participants ($M = .63$, $SD = .39$). Results also indicated a significant effect of diagrams, $F(1, 66) = 4.1, p = .047, \eta^2_p = .06$; students learning from text with diagrams improved their mental models ($M = .79, SD = .33$) more than students using text alone ($M = .48, SD = .43$). The quadratic effect and the interaction between knowledge level and the quadratic effect were not significant; thus, mental model improvement was not significantly better when using simplified diagrams, nor was there an interaction of diagram type and background knowledge.

It should be noted that, unlike the premodel drawing task, the postmodel drawing task did show differences in the number of requests for clarification made by the experimenter. First, experimenter prompts differed by experimental condition, $F(2, 63)^{1} = 3.7, p < .05, \eta^2_p = .11$. Tukey–Kramer post hoc analyses demonstrated that participants using detailed diagrams required more requests for clarification than participants using simplified diagrams ($M_{\text{Diff}} = 2.7, \text{Critical}_{\text{Diff}} = 1.5$) or text only ($M_{\text{Diff}} = 1.8, \text{Critical}_{\text{Diff}} = 1.5$), where $M_{\text{Diff}}$ refers to the mean difference between groups and Critical$_{\text{Diff}}$ refers to the levels of critical difference. In addition, a difference was found in the number of experimenter prompts based on postmodel type, $F(3, 62) = 10.7, p < .001, \eta^2_p = .27$. Tukey–Kramer post hoc analyses showed that students who generated a fully correct double loop 2 postmodel needed significantly fewer prompts than students who did not form the correct mental model ($M_{\text{Diff}} = -5.2, \text{Critical}_{\text{Diff}} = 1.8$). The interaction between experimental condition and postmodel was not significant ($F < 1$). Together, these results foreshadow results from the other learning measures tested in Experiment 1: Although participants using detailed diagrams did improve their mental models, they often appeared to be confused about what they had learned and about how to revise their mental models. As may be expected, students who learned the correct model of the heart and circulatory system were better able to communicate their understanding of the domain than students who failed to form a correct understanding of the domain. Neither of these results is particularly surprising, but both results support the idea that diagram representation influences learning and that students who fail to learn a fully correct model of the domain often experience some confusion about their own mental models.

The finding that diagrams assist mental model development is consistent with the multimedia effect (e.g., Mayer, 2001; Winn, 1987); current results demonstrate that diagrams presented with text support development of students’ mental models of the domain. In addition, the finding that greater prior domain knowledge supports learning is consistent with previous research in text comprehension (see Kintsch, 1998); it is not unexpected that higher knowledge participants would be better able to refine their mental models than participants with low prior knowledge.

Pre- to posttest improvement in general knowledge. Analysis of pre- to posttest improvement on general knowledge demonstrated a significant quadratic effect, $F(1, 66) = 6.09, p = .017, \eta^2_p = .10$, such that, overall, participants using simplified diagrams outperformed those using detailed diagrams and text alone (see Figure 3). No other significant effects or interactions were identified. There were no differences in pretest scores by experimental condition ($F < 1$). Thus, students using simplified diagrams gained the most knowledge about general information concerning the heart and circulatory system.

---

1 Because of equipment failure during recording, one premodel and one postmodel protocol could not be analyzed.
Memory questions about the text. Results demonstrated a significant effect of prior knowledge, $F(1, 66) = 19.4, p < .0001, \eta^2_p = .24$; participants with high prior knowledge remembered more about the text than participants with low prior knowledge. Results also showed a significant quadratic effect, $F(1, 66) = 4.36, p = .041, \eta^2_p = .07$, such that, overall, participants using simplified diagrams outperformed those using detailed diagrams and text alone (see Figure 3); it is interesting to note that the size of this effect is intensified when considering only the lower knowledge participants, $F(1, 50) = 7.92, p = .007, \eta^2_p = .14$. No other significant effects or interactions were identified.

The memory results are similar to those found for general knowledge questions. Simplified diagrams appear to benefit participants differently than do detailed diagrams or text alone. Simplified diagrams best supported participants’ learning of factual information as measured by pre- to posttest improvement on general domain knowledge and by memory for specific text information.

Inference questions related to the text. There were no experimental condition effects on inference question scores. Participants using diagrams did not outperform students using text only ($F < 1$), nor was there an advantage for simplified diagrams ($F < 1$) or an interaction between condition and knowledge level, $F(1, 66) = 1.95, p > .16$. However, there was a main effect of prior knowledge such that high-knowledge participants outperformed lower knowledge participants, $F(1, 66) = 20.04, p < .0001, \eta^2_p = .24$. Overall, scores on the inference questions were quite low and likely reflect a floor effect. As may be expected by the apparent

![Figure 3. Overall summary of learning outcomes: Mean performance of each experimental condition on learning measures in Experiment 1 (left column) and in Experiment 2 (right column).](image-url)
Difficulties of these questions, high-knowledge learners were better able to apply their knowledge to new situations, but even these students had difficulty with the inference questions. In general, the inference questions required students to reason about both the structure and function of the heart in order to arrive at a correct solution. It is possible that questions requiring fewer inferences with more direct transfer opportunities would have resulted in better comprehension data.

Experiment 1 Summary

Diagrams help, but simplified diagrams help most. Experiment 1 demonstrated that both types of diagrams were useful tools for mental model development but that simplified diagrams best facilitated factual learning in the domain. Somewhat surprisingly, focused tests of learning outcomes showed that learners using detailed diagrams were statistically comparable with learners using only text when improvement in general knowledge and text memory was tested. No interaction between learner background knowledge and diagram effectiveness was identified. Contrary to expectations, high-knowledge learners did not benefit more from detailed diagrams. Results did show a strong multimedia effect when mental model development is considered alone, but, overall, learners appeared to be best supported by simplified diagrams when all learning outcomes were considered together.

The current results are quite consistent with previous research showing that diagrams can improve memory and understanding (e.g., Mayer, 1989; Mayer & Anderson, 1992; Mayer & Gallini, 1990). However, they also highlight an important consideration: The effectiveness of diagrams can be influenced by their representation. Diagrams appear to be increasingly effective when they are designed to support learners by clearly depicting key functional relationships, even when this support is accomplished by abstraction and simplification.

Evidence for informational equivalence. An alternative explanation for the superiority of simplified diagrams in Experiment 1 may be that participants were simply unable to interpret the structural complexity of the detailed diagrams. If this were the case, then detailed diagrams could not be expected to support learning or memory better than text alone. To test this possibility, I asked 6 participants from the University of Colorado at Boulder—all with low prior knowledge of the heart and circulatory system—to fully label and annotate a detailed diagram of the heart and circulatory system while they studied the experimental materials depicting detailed diagrams. All 6 students produced complete and entirely accurate labels and annotations that indicated a fully correct model of the heart and circulatory system.

These data demonstrate that participants can interpret detailed diagrams accurately. When the experimental task required them to attend to the detailed diagrams and to continually update an externally represented version of the detailed diagram, they had no trouble doing so. Thus, the results from Experiment 1 cannot be attributed to informational poverty or undue complexity in the detailed diagrams. The benefit of simplified diagrams may be that they support comprehension processes during learning and result in better integration between the verbal (text) and visual (diagram) representations. Experiment 2 explores the underlying reasons for the learning benefits associated with diagrams—particularly simplified diagrams—found in Experiment 1.

Experiment 2

Evidence from Experiment 1 demonstrated that diagrams promoted improved understanding of domain information in the form of more accurate mental models of the heart and circulatory system. Combined with previous evidence that the addition of diagrams and illustrations to learning materials can support deep understanding (e.g., Mayer & Anderson, 1992; Mayer & Gallini, 1990), experimental evidence suggests that diagrams not only support memory for information but also promote situation model construction during learning. The situation model is one of three levels of text representation identified by van Dijk and Kintsch (1983); the other two levels are the surface level and the textbase. The surface level and the textbase correspond to memory for information. Surface information corresponds to the exact words and phrases of a text, whereas the textbase is more closely associated with what most people consider as memory for information; the textbase contains relations and ideas derived directly from a text but not necessarily in the exact surface representation provided in the text. In contrast, the situation model extends beyond the information given in the text; the situation model reflects a more flexible and deeper representation that forms when current information is interpreted and integrated with prior knowledge. A good situation model of a text allows the learner to make inferences and to apply knowledge to other situations (Kintsch, 1998).

Thus, it was hypothesized that diagrams would support the generation of inferences and the integration of information during learning, as would be expected if diagrams support construction of a well-developed situation model during learning. Because simplified diagrams were designed to highlight the component relationships in the heart necessary to understanding the heart and circulatory system, it was hypothesized that simplified diagrams would be most effective in promoting inferences that integrate the to-be-learned information.

Using a self-explanation methodology, Experiment 2 explored the types of comprehension processes in which students engaged when learning with diagrams and text. Experiment 2 also investigated whether diagram representation (simple vs. detailed) would influence comprehension processes during learning.

Method

Participants

Participants were 34 undergraduates at the University of Colorado at Boulder. All were native English speakers and received course credit for their participation. Because random sampling produced very few high-knowledge participants in this experiment, only data from 25 low-knowledge (single loop premodels) participants were analyzed.

Materials and Procedure

The materials and procedure in Experiment 2 were the same as Experiment 1 with one exception: Experiment 2 participants self-explained during learning. All participants were given written instructions about self-explaining; the experimenter orally reviewed the instructions with each student and asked for questions before students began the learning phase. The addition of self-explanation during the learning phase increased its length by 10–20 min from Experiment 1.

Although students self-explained spontaneously, the experimenter did prompt students to continue explaining when their comments were vague, incomplete, or when the student paused for more than about 10 s during the
protocol. During the first two pages of material, the experimenter provided all participants with feedback about their self-explanations. Experimenter prompts to students during learning did not vary by experimental condition when analyzing the total number of prompts given, $F(2, 22) = 2.0, p > .16$, or the overall number of pages prompted, $F(2, 22) = 1.1, p > .35$. All comments were recorded on audiotape.

**Analysis of Self-Explanations**

Students’ comments were transcribed and then separated into a series of complex propositions that correspond roughly to an idea unit (see Kintsch, 1998, for more information on complex propositions). Sentences uttered by a participant often reflected propositions of more than one type (e.g., a paraphrase and then an inference), so scoring complex propositions permitted accurate analysis of the data. Appendix B shows 1 participant’s comments about a section of text and the transformation from the original transcription to coded complex propositions of the student’s statements. Complex propositions of each participant’s self-explanations were categorized into one of four general categories, listed below. Sample self-explanations from each category are shown in Appendix C.

**Paraphrases.** These self-explanations reflected restatements of text displayed to students at the time of utterance. Paraphrases were not necessarily exact replications of the text; a complex proposition conveying an idea given in the current text was scored as a paraphrase.

**Elaborations.** These self-explanations reflected connections to a participant’s background knowledge that could not be inferred solely from the learning materials.

**Monitoring statements.** These self-explanations were metacognitive statements that reflected checks on understanding, confusion, or questions about the materials.

**Self-explanation inferences.** Self-explanation inferences were statements that reflected knowledge or inferences generated during self-explaining that worked with the learning materials but that went beyond learning material displayed at the time of utterance. A piece of new information inferred from displayed materials would be coded as a self-explanation inference as would a statement that reflected integration of current material with previously presented information. Because inferences that integrate information are particularly important to situation model development, these inferences were separately coded and analyzed. Any inference that integrated previously stated information with the current information was coded as an integration inference.

**Errors.** Complex propositions, in any category, that reflected clearly incorrect content also were scored as errors. Ambiguous or partially correct statements were not scored as errors.

**Interrater reliability.** Two raters scored the complex propositions generated by participants. After a brief (20-min) training session, raters made an initial scoring pass at three protocols from each experimental condition. Raters rescored the data after resolving questions about category criteria. Interrater correlations for statements in the major scored categories were quite high, ranging from a minimum of .70 to a maximum of .99 ($p < .001$).

Table 2

<table>
<thead>
<tr>
<th>Statement type</th>
<th>Interrater correlation</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paraphrases</td>
<td>.92</td>
<td>.0001</td>
</tr>
<tr>
<td>Elaborations</td>
<td>.70</td>
<td>.0369</td>
</tr>
<tr>
<td>Monitoring</td>
<td>.99</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Self-explanation inferences</td>
<td>.90</td>
<td>.0003</td>
</tr>
<tr>
<td>Integration inferences</td>
<td>.81</td>
<td>.0062</td>
</tr>
<tr>
<td>Errors</td>
<td>.91</td>
<td>.0002</td>
</tr>
</tbody>
</table>

Raters rescored the data after resolving questions about category criteria. Inference integration was quite high ($\eta^2 = .94$, $\eta^2_{p_{o.s.m.}} = .96$); overall, three initial disagreements were resolved using discussion until 100% agreement was reached. As in Experiment 1, both types of diagrams did support mental model development, $F(1, 24) = 8.97, p = .0067$, $\eta^2 = .29$. However, the quadratic effect testing the mean of the simplified diagram group against those of the detailed diagram and text-only groups was also significant, $F(1, 24) = 7.19, p = .014$, $\eta^2 = .25$; this quadratic pattern is evident in Figure 3.

Learning Outcomes

Figure 3 provides a visual summary of the learning outcomes analyzed in these experiments. As can be seen in this figure, the pattern of results identified for Experiment 2 learning outcomes (right-hand column) was very consistent with Experiment 1 (left-hand column). Low-knowledge participants again appear to be at a performance advantage when using the simplified diagrams during learning.

**Mental model improvement.** Interrater reliability for mental model identification was quite high ($\kappa_{w_{p_{o.s.m.}}} = .94$, $\kappa_{w_{p_{o.s.m.}}} = .96$); overall, three initial disagreements were resolved using discussion until 100% agreement was reached. As in Experiment 1, both types of diagrams did support mental model development, $F(1, 24) = 8.97, p = .0067$, $\eta^2 = .29$. However, the quadratic effect testing the mean of the simplified diagram group against those of the detailed diagram and text-only groups was also significant, $F(1, 24) = 7.19, p = .014$, $\eta^2 = .25$; this quadratic pattern is evident in Figure 3.

There were no differences in the number of experimenter requests for clarification between experimental groups during the predraw, $F(2, 22) = 1.65, p > .21$, or the postdraw, $F(2, 21) = 1.62, p > .22$; because of recording equipment error, one postmodel could not be analyzed. Experimenter requests for clarification during the postdraw also did not vary by postmodel ($F < 1$), likely a consequence of requiring self-explanation during the learning phase.

---

2 It should be noted that a linear contrast code would be orthogonal to the quadratic test. However, a linear effect is not theoretically meaningful in this experiment; the contrast codes chosen reflect the substantive hypotheses of this experiment.
It is interesting to note that results continued to demonstrate differences between experimental conditions even though all participants were engaged in active processing during learning (self-explaining). It might be argued that diagrams are inherently interesting and increase the amount of time students spend studying the materials or self-explaining. However, there were no differences in the amount of time spent studying materials when comparing diagram conditions with text only, \( F(1, 24) = 1.2, p > .28 \), or simplified diagrams with the other conditions \((F < 1)\); overall, students spent an average of 29 min (minimum = 19, maximum = 43) studying the experimental materials. Thus, diagram benefits cannot be attributed to a simple increase in time on task and may indicate a possible gain in efficiency. This possibility is consistent with recent work by Ainsworth and Loizou (2003). In Ainsworth and Loizou’s study, students using diagrams generated more self-explanations than students using text only even though the diagrams participants uttered fewer words and spent significantly less time learning the materials. Thus, the possibility that diagrams may spur more efficient learning is worth further exploration in future research.

Pre- to posttest of general (factual) knowledge. Pretest scores were statistically equivalent for all experimental conditions \((F < 1)\). As in Experiment 1, the quadratic effect testing the mean of the simplified diagrams group against the means of the detailed diagrams and text-only groups was significant, \( F(1, 24) = 6.16, p = .021, \eta^2_p = .22 \), and can be seen in Figure 3. The test of an overall diagram effect was not significant, \( F(1, 24) = 1.0, p > .32 \). These results reinforce the advantage of simplified diagrams, consistent with findings from Experiment 1.

Memory questions about the text. Again, the quadratic effect testing the superiority of the simplified diagrams group was significant, \( F(1, 24) = 6.50, p = .018, \eta^2_p = .23, \) and the test of an overall diagram effect was not \((F < 1)\). As seen in Figure 3, a quadratic pattern is evident in nearly every learning outcome tested with lower knowledge participants.

In both Experiments 1 and 2, simplified diagrams facilitated specific memory for the learned information, whereas the detailed diagrams failed to support memory to a greater extent than text alone. It is noteworthy that a consistent benefit for simplified diagrams has been found across experiments for multiple measures.

Inference questions related to the text. There were no experimental condition effects for the inference questions in this experiment \((F < 1)\). As in Experiment 1, low-knowledge participants performed near floor on the inference questions. As discussed in Experiment 1, poor performance may reflect that these inference questions were too detailed and required too many related inferences from participants.

Self-Explanations

The patterns of self-explanations generated by participants overall and in each experimental condition are presented in Table 4. Other statements versus self-explanation inferences. The predicted contrasts (the test of overall diagram benefit and the quadratic effect testing superiority of simplified diagrams, respectively) were used to test for potential differences in comprehension processes based on the a priori hypotheses. No differences were found for paraphrases \((F < 1), F(1, 24) = 1.4, p > .25 \), elaborations \((F < 1), F(1, 24) = 1.8, p > .19 \), or monitoring statements \((F < 1)\) generated by participants. For simplicity, these statements are grouped together as Other Statements in Figure 4. In general, participants made a large number of statements other than self-explanation inferences during the protocols, but the addition of diagrams to text did not influence generation of these statements.

The mean number of self-explanation inferences generated by participants in each experimental condition also is shown in Figure 4. Participants using simplified diagrams did not generate significantly more inferences than the other groups, \( F(1, 24) = 4.3, p > .05 \), but diagrams overall did promote a significant increase in inference generation compared with text only, \( F(1, 24) = 14.2, p = .001, \eta^2_p = .39 \). Thus, a multimedia effect was identified for inference generation.

Comparing the number of self-explanation inferences and other statements generated by participants in the experimental conditions reveals a striking pattern. The addition of diagrams to text had a very strong effect on the number of inferences that participants generated but did not affect other comprehension processes in which students engaged while self-explaining. This result supports recent findings by Ainsworth and Loizou (2003); these authors investigated diagram effects on self-explanations and found that participants using diagrams generated significantly more self-explanations but an equivalent number of monitoring statements to students who used text only.

Erroneous self-explanation inferences. Although participants in the text-only condition generated fewer inferences, they were more than twice as likely to be in error when making a self-explanation inference compared with participants in the diagram conditions, \( F(1, 24) = 5.8, p < .025, \eta^2_p = .21 \). However, as
assessed by the quadratic contrast, simplified diagrams were not 
most effective at preventing inference errors, $F(1, 24) = 1.5, p > .23$.

Thus far, the results demonstrate a strong multimedia effect in 
comprehension processes: Both types of diagrams encourage the 
generation of correct self-explanation inferences when participants 
self-explain a text. But how do diagrams affect specific inferences 
associated with situation model development in the domain?

Integration self-explanation inferences. Inferences reflecting 
the integration of learned material may be particularly important to 
the development of an accurate mental model. As seen in Figure 5, 
participants using diagrams generated more integration inferences 
than students using text only, $F(1, 24) = 5.3, p = .031, \eta_p^2 = .20$, 
but the effect is stronger when analyzing the potential superiority 
of simplified diagrams compared with the other conditions, $F(1, 
24) = 8.9, p < .007, \eta_p^2 = .29$. Thus, the pattern of integration 
inferences indicated that simplified diagrams support integration 
inferences to a greater extent than do detailed diagrams or text-
only materials. Further, generation of integration inferences is tied 
to developing a correct mental model of the domain; learners who 
formed the correct double loop 2 model of the heart and circulatory 
system generated significantly more integration inferences than 
learners who formed an incorrect model, $F(1, 23) = 6.08, p < .022, \eta_p^2 = .21$.

The advantage of simplified diagrams in promoting the gener-
ation of integration inferences may help explain the advantage of 
simplified diagrams identified in these experiments. Integration of 
information is a key component to developing a complete and 
accurate representation of the to-be-learned material; thus, it may 
not be surprising that learners using simplified diagrams demon-
strated increased generation of integration inferences in addition to 
a performance advantage on learning outcome measures.

General Discussion

In Experiments 1 and 2, participants who used diagrams with 
text were best able to improve their mental models, whereas 
participants in the text-only condition improved their mental mod-
els least. This finding demonstrates clear and compelling evidence 
for the multimedia effect. However, these experiments also iden-
tify a learning advantage when diagrams have been carefully 
designed to highlight the representation of critical relationships of 
the domain information. Consistent with the mental model results, 
measures testing improvement in general knowledge and memory 
repeatedly identified the superiority of simplified diagrams over

<table>
<thead>
<tr>
<th>Proposition type</th>
<th>Overall</th>
<th>Text only</th>
<th>Simplified diagrams</th>
<th>Detailed diagrams</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total statements</td>
<td>255</td>
<td>202</td>
<td>278</td>
<td>270</td>
</tr>
<tr>
<td>Paraphrases</td>
<td>90</td>
<td>85</td>
<td>99</td>
<td>83</td>
</tr>
<tr>
<td>Elaboration</td>
<td>23</td>
<td>20</td>
<td>29</td>
<td>19</td>
</tr>
<tr>
<td>Monitoring</td>
<td>44</td>
<td>38</td>
<td>38</td>
<td>16</td>
</tr>
<tr>
<td>Self-explanation inferences</td>
<td>87</td>
<td>46</td>
<td>102</td>
<td>103</td>
</tr>
<tr>
<td>Integration inferences</td>
<td>26</td>
<td>17</td>
<td>34</td>
<td>23</td>
</tr>
<tr>
<td>Overall errors (%)</td>
<td>5</td>
<td>7</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Inference errors (%)</td>
<td>13</td>
<td>22</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

Note. For simplicity, all numbers are rounded to the nearest integer.
detailed diagrams and text-only materials. Why were simplified diagrams so powerful? Self-explaining data from Experiment 2 may help provide an interpretation.

Comprehension Processes: Self-Explanations

Diagrams Encourage Correct Self-Explanation Inferences

The self-explaining data from Experiment 2 offer a striking conclusion: The influence of diagrams on comprehension processes was specific to inference generation. Diagrams increased the frequency with which learners generated inferences but did not influence production of other statements.

Further, the current results demonstrated that diagrams supported learners in generating correct inferences. Comprehension research has shown that adequate domain knowledge is important to learning in large part because it facilitates generation of correct inferences. Moravcsik and Kintsch (1993) found that although higher and lower knowledge readers generated approximately equal numbers of inferences during learning, most of the inferences generated by lower knowledge readers were in error. Thus, the assistance of diagrams in supporting correct inference generation with lower knowledge learners may place them at a significant comprehension advantage, supporting the multimedia effect that has been repeatedly identified in previous research (e.g., Mayer, 2001).

Diagrams Encourage Different Types of Inferences

Diagrams also may influence the type of inferences generated by participants. Results indicated that participants using simplified diagrams generated more integration inferences than participants using detailed diagrams or text-only materials. To the extent that integration of previous and current information is necessary to form an adequate textbase and situation model (e.g., Kintsch, 1998), heightened support of integration inferences may explain why participants using simplified diagrams outperformed participants using detailed diagrams in nearly every outcome measure.

It should be noted that participants in the diagram conditions may have had additional support for integrating information because the diagrams essentially provide a summary of previously learned information whenever a diagram is present during learning. As a visual summary, diagrams may prompt students to integrate information or may provide important cues for recall of previous information. However, diagrams were not equally effective in supporting integration of information; thus, the effectiveness of such visual summary support may depend on the success of the representation in highlighting crucial relationships necessary to understand the text. Future work should address whether well-designed text summaries may prompt integration benefits similar to those found with diagrams.

Self-Explaining With Diagrams: Representation for Strategic Support

Although all participants made a large number of statements when self-explaining, only participants using diagrams increased their generation of inferences. Students using the text-only materials generated a large number of statements when self-explaining but demonstrated markedly less successful learning outcomes. Why does the same type of activity result in this processing difference? Previous research has found that readers who actively work with text materials are more likely to remember the text and to be able to solve comprehension problems after learning (Chi et al., 1994; Goldman & Saul, 1990; McNamara & Kintsch, 1996). However, current results suggest that not all types of active processing are equivalent. Certain types of active processing—specifically, comprehension processes associated with deeper learning—lead to greater performance gains than others.

The current results demonstrate that learners using diagrams are supported in generating inferences and that simplified diagrams specifically support the integration of information necessary for deep learning. Thus, participants using diagrams engage in more useful comprehension processes more frequently. In essence, diagrams provide a type of cognitive support that may help explain why participants who received diagrams (especially participants receiving simplified diagrams) demonstrated greater learning even though they were not found to spend significantly more time on task than participants in the text-only condition.

Designing Diagrams for Effective Learning: Attending to Cognitive Support

Overall, results suggest that when a simplified diagram guides the learner appropriately, too much diagram complexity is unnecessary and likely unhelpful. However, one must be careful not to misinterpret this finding. Results from self-explaining suggested that the simplified diagrams were beneficial because they promoted generation of important inferences during learning. Thus, diagrams may be most useful when they have been designed to highlight the essential relationships necessary to understand the situation described in the learning materials. In the current experiments, simplified diagrams highlighted essential domain relationships and encouraged the cognitive processes necessary for deep learning. In other situations, simplified diagrams may not demonstrate strong benefits unless they can effectively guide learners in inference generation or other important comprehension processes.

Taken as a whole, the results from these experiments suggest that design and evaluation of multimedia should include careful consideration of cognitive support. The idea of cognitive support can be seen as a general principle that supports existing multimedia principles; it proposes that visual materials will be most effective when they are designed to support cognitive processing of critical information. As demonstrated in this research, it is not enough to simply encourage active engagement with learning materials; rather, good multimedia materials support the learner’s effective performance of essential cognitive processes. Effective impact on learning requires that representations can successfully support the cognitive processes necessary for deep comprehension.

It appears that the power of visual representations lies in their ability to support the learner’s strategic use of comprehension processes during learning rather than in the specific characteristics of the media itself. Accordingly, developers and researchers are faced with the difficult task of determining how multimedia materials can most effectively support student performance of the processes critical to learning. Researchers need to know more about the ways that experts learn with diagrams and multimedia, and they must determine how to use visual representations to support these processes. But, as a general cognitive approach to multimedia learning, attending to the cognitive support of diagrams represents another step toward the continued refinement of effective principles for multimedia design.
References


American Heart Association. (1988). *How your heart works (lower elementary): The heart.* (Available from the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231–4596)


(Appendices follow)
Appendix A

Sample Memory and Inference Questions

Sample General Knowledge Questions (in order of difficulty)

- How many chambers are there in the heart?
- Blood returning from the body enters which chamber of the heart first?
- How many continuous, closed circuits of blood are there from the heart?
- What is an atrium?
- What is a ventricle?

Sample Memory Questions About the Text

- How many valves are there in the heart?
- What is the average heart rate for a child?
- On average, how much does an adult human heart weigh?
- The sound of the heartbeat is often characterized as “lub-dub.” To what activities do the “lub” and the “dub” correspond?

Sample Inference Questions Related to the Text

- Carbon monoxide (CO) binds with the hemoglobin in red blood cells better than it does with either oxygen or carbon dioxide. If someone has inhaled a great deal of CO, what type of gas exchange will take place in the lungs of that person?
- What would happen if the valves leading out of the ventricles didn’t close properly?
- Assuming that the strength and timing of contractions would remain constant, what effect if any would result from separating the left and right halves of the heart? Why?
- Imagine a disease that causes the walls of all blood vessels in the body to grow dramatically thicker, effectively narrowing the width of all the blood vessels in the body. What effect do you think this would have on the heart and the circulatory system?

Appendix B

Sample Complex Propositions

Text From Learning Materials

“As the blood flows through the capillaries in the body, carrying its supply of oxygen, it also collects carbon dioxide. The blood that empties into the right atrium is dark colored. It has picked up carbon dioxide from the body cells” (Silverstein & Silverstein, 1983, p. 13).

Transcription of Participant #2014 Statements

So, when the blood flows through the capillaries, the blood rich, er, the oxygen rich blood flows through the capillaries and collects carbon dioxide. And . . . I guess it turns the blood dark. Well, it says that the blood that empties into the right atrium is dark, because it picks up the carbon dioxide from the blood cells. So, I’m guessing it turns it dark. I don’t know, I guess the blood is dark because of the carbon dioxide and the oxygen probably, like, enriches the red color. Makes the red color.

Table B1

Complex Propositions and Categorization

<table>
<thead>
<tr>
<th>Complex proposition</th>
<th>Coded category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Blood flows through the capillaries</td>
<td>Paraphrase</td>
</tr>
<tr>
<td>2. The oxygen-rich blood flows through the capillaries</td>
<td>Paraphrase</td>
</tr>
<tr>
<td>3. [2] and collects carbon dioxide</td>
<td>Paraphrase</td>
</tr>
<tr>
<td>4. Carbon dioxide turns the blood dark</td>
<td>Inference</td>
</tr>
<tr>
<td>6. Blood that empties into the right atrium is dark</td>
<td>Paraphrase</td>
</tr>
<tr>
<td>8. [6] Because blood picks up the carbon dioxide from the cells</td>
<td>Inference</td>
</tr>
<tr>
<td>9. [6, 8] so I’m guessing carbon dioxide turns blood dark</td>
<td>Inference</td>
</tr>
<tr>
<td>10. I don’t know</td>
<td>Monitoring</td>
</tr>
<tr>
<td>11. The blood is dark because of the carbon dioxide</td>
<td>Inference</td>
</tr>
<tr>
<td>13. Oxygen probably enriches the red color of the blood</td>
<td>Inference</td>
</tr>
<tr>
<td>14. Oxygen makes the red color of the blood</td>
<td>Inference</td>
</tr>
</tbody>
</table>

Note. Bracketed information refers to previously numbered complex propositions.
Appendix C

Sample Self-Explanations

Original Text

“The blood does both these things. It brings oxygen to the body cells and takes away their carbon dioxide. These gases pass easily back and forth through the thin walls of the tiny capillaries” (Silverstein & Silverstein, 1983, pp. 11–12).

Sample Paraphrases

Student 2008: Blood brings oxygen to the body cells. Blood takes away the cells’ carbon dioxide.
Student 2013: Carbon dioxide gets into the tiny capillaries.
Student 2014: The gases can go through the wall of the capillaries.

Sample Elaborations

Student 2021: There’s blood in muscle tissue.
Student 2024: [Carbon dioxide goes back into the blood] so carbon dioxide can be expelled from the body.
Student 2029: Carbon dioxide is when you exhale.

Sample Monitoring Statements

Student 2007: I remember that [the gases flow back and forth from the capillaries].
Student 2019: I knew capillaries had something to do with giving oxygen to the body cells.
Student 2021: I don’t know why the gasses gotta get out of the capillaries.

Sample Inferences

Student 2007: [Oxygen needs to get replenished, so] blood goes back into the heart and in the lungs.
Student 2013: Carbon dioxide gets pumped back into the heart. Then carbon dioxide is taken to the lungs.
Student 2020: [After blood gets oxygen in the lungs] oxygen is carried to the heart.
Student 2010: We have the two separate circles of carbon dioxide and oxygen.
Student 2030: We would not be able to live if the heart didn’t do one of these things.

Integration Inferences

Student 2008: When you eat and combine oxygen, that creates carbon dioxide. The little capillaries [through which gases pass back and forth] are furthest away from the heart.
Student 2015: If CO2 stayed the cells would be poisoned.
Student 2017: Capillaries gather carbon dioxide so it won’t poison the cells.

C1 Content in brackets was represented in other complex propositions generated for the current text by the participant.

Received January 13, 2004
Revision received July 15, 2005
Accepted September 8, 2005