Learning With Analogy and Elaborative Interrogation

Mark A. McDaniel  
University of New Mexico

Carol M. Donnelly  
Purdue University

Previous work has found that embedding analogy in a text improves accurate inferencing, but at the expense of factual learning of newly learned scientific concepts. This study explored the possibility of eliminating the decrease in factual learning by combining analogy with key-word highlighting, pictorial schematics, or elaborative interrogation. Schematics had no effect on either factual or inference learning. Combining key-word highlighting with analogy increased factual learning to levels comparable with those found in the literal-text conditions. Elaborative interrogation produced robust gains in both factual and inference learning, regardless of whether the technique was combined with analogy. These results represent an extension of the situations in which elaborative interrogation produces potent learning benefits and emphasizes its potential over alternative instructional methods.

What is the most effective way to teach learners scientific concepts? Much of what we know about this issue is based on research that focuses on acquisition of factual information (e.g., Martin & Pressley, 1991). Certainly this is one aspect of learning about scientific concepts, but it is by no means the only aspect of acquiring scientific concepts. As many educational psychologists have noted, a main objective in learning is to be able to go beyond the particular facts stated and use the conceptual information in new ways and in new contexts. That is, it is important to teach about the scientific concept in such a way as to facilitate inferential thinking (e.g., Mayer & Gallini, 1990). (Indeed, it may be that this kind of deep understanding of concepts is a prerequisite for creative thinking; see Donnelly, 1995.) In this article, one of our main objectives is to explore and identify learning techniques that foster inferential thinking with newly acquired scientific concepts.

Donnelly and McDaniel (1993) explored the value of analogy for learning elementary concepts in the domains of astronomy, biology, and physics. They compared students’ learning when concepts were described through an analogy with learning when concepts were stated literally. Learning was measured through a series of multiple-choice questions that tapped either stated facts about the concept (factual-level questions) or appropriate inferences that could be drawn from the concept (inference-level questions). Embedding concepts in an analogical framework consistently promoted inference-level performance compared with when the concepts were presented with literal statements alone. This benefit, however, was at the cost of learning factual information about the scientific concepts. Participants in the analogy condition performed less well when tested for factual information than those provided with the literal rendition of the scientific concepts.

Although it is encouraging that analogy facilitated inferential learning, we would argue that analogy, as used by Donnelly and McDaniel (1993), is not optimal because it does not enable comprehensive acquisition of scientific concepts (i.e., it does not improve learning of both factual and inferential knowledge). Our theoretical interpretation is that the core benefit of using analogy is that the learner can apply a familiar structure or mental model (activated from the source) to help ascertain the fundamental and relevant relations in the target domain (Gentner, 1983; Shustack & Anderson, 1979). That is, our working hypothesis is that embedding analogy in a text affords relational processing (a type of elaboration that relates the components of the text together). Indeed, the relational structure of the source appeared to be so compelling in the Donnelly and McDaniel study (Experiment 4) that a complete absence of specific facts concerning the target domain did not disrupt the inferencing advantage conferred by the source. This result seems related to other work in text memory, showing that if the text material itself invites relational processing, then the learner’s processing of specific and detailed facts can be obscured (Einstein, McDaniel, Bowers, & Stevens, 1984; see McDaniel & Einstein, 1989, for a review).

In a similar vein, we suggest that the high-level relational processing presumably induced by analogy reduces attention to factual details. This reasoning leads to several predictions that we test in the following experiments. First, and more important from an applied standpoint, an apt analogy combined with an appropriate study adjunct could provide more comprehensive learning than analogy alone could provide. Second, our distinction between higher order relational processing and processing of specific facts allows a
priori specification of what kinds of study adjuncts will effectively combine with analogy. Specifically, a study manipulation that encourages students to focus on detailed facts should invite cognitive processes complementary to the high-level relational processing promoted by the analogy. In contrast, a study manipulation that encourages students to focus on high-level relations in the text would be redundant with the processing prompted by analogy and, consequently, should not benefit learning (see Einstein, McDaniel, Owen, & Coté, 1990). We tested these predictions in Experiment 1.

Third, reasoning along the same lines as above, in Experiments 1 and 2 we explored the possibility that an appropriate study adjunct could be selected for the literal presentation of scientific concepts to facilitate inferencing and thereby provide comprehensive learning of the concepts.

Experiment 1

The primary focus of this experiment was to explore learning adjuncts that could be combined with analogically based text to foster more comprehensive learning than analogy alone could provide. Specifically, we attempted to develop learning adjuncts to override the diminished factual learning associated with analogy-based text (relative to a text without analogy, labeled a literal text). One of the novel features of this effort is that we based selection of the learning adjuncts on a prior theoretical analysis, an analysis that anticipated the particular learning adjuncts that would successfully enhance learning for the analogy-based text, as well as those that would not.

The theoretical orientation that we adopted is the material-appropriate processing framework (MAP; McDaniel & Einstein, 1989). This framework suggests that for a learning adjunct to be effective, the type of processing evoked by the learning adjunct should complement the type of processing encouraged by the text material itself. For example, if the textual material intrinsically encourages relational processing (elaboration that organizes the components of the text), then learning is substantially enhanced with a learning adjunct that emphasizes processing of individual facts (provided this learning adjunct does not disrupt the organization of the material; see Guynn, Einstein, & Hunt, 1992) but not with a learning adjunct that emphasizes relational processing (Einstein et al., 1990). Along the same lines, we anticipated that successful learning adjuncts for the analogy-based texts would be those that focused the learner on detailed facts about the scientific concepts described in the texts. In contrast, we expected that learning adjuncts that encouraged attention to the relational structure of the facts would not enhance learning for the analogy-based text. These two main predictions are based on the assumption that analogy itself helps the student focus on the important relations implicit in a scientific concept and on the MAP idea that only learning adjuncts that complement the processing invoked by the text itself will enhance learning.

To encourage more processing of individual facts, we highlighted key words in the text and instructed the reader to focus on the key words. This key-word adjunct was applied to both an analogy-based text and the corresponding literal text. On the basis of the ideas just developed, we expected that the key-word analogy condition would show improved fact learning over the analogy-alone condition and as good or better fact learning than either the key-word or basic literal conditions. In contrast, we expected no improvement due to the key-word adjunct for the literal-text condition. The literal condition appears already to invite processing of individual facts (Donnelly & McDaniel, 1993). That is, in the absence of a relational structure (presumably provided through analogy), learners may focus on encoding the individual facts or propositions presented in the passage (cf. Graf & Levy, 1984; Kintsch & Young, 1984). Again, the idea here is that the key-word adjunct and the literal text invite similar types of processing, and according to MAP such redundant processing does not enhance learning. Finally, the key-word adjunct was not expected to aid performance on inference-level test questions.

We implemented two additional adjuncts to encourage relational processing. We added these adjuncts to test the prediction that requiring readers to focus more attention on the content of the texts per se is not sufficient to produce more comprehensive learning (i.e., more factual learning) in the analogy-based text. One relational adjunct was a labeled schematic that pictorially displayed how certain key concepts of the scientific construct are related (see the Appendix for an example). We selected this adjunct as a strong test of the ideas developed here. Labeled schematics have previously been shown to improve students’ learning performance on certain kinds of technical texts (Mayer & Gallini, 1990). Therefore, the anticipated failure of this adjunct to aid learning for analogy-based text could not be attributed to the fact that it is a nonpotent learning adjunct in general.

Still, it might be argued that an experimenter-provided schematic does not require the learner to exert as much cognitive effort (see Tyler, Hertel, McCallum, & Ellis, 1979) as is expended when students are instructed to focus on underlined material (i.e., the key-word adjunct). On this argument, nonsignificant effects of schematic adjuncts would be due to insufficient cognitive effort, not to an inappropriate type of processing associated with the adjunct. Accordingly, we included a second relational adjunct that required the reader’s involvement: Students were instructed to generate their own labeled schematic. Although we did not directly gauge the cognitive effort associated with each adjunct, it seems unlikely that generating a labeled schematic based on the contents of the passage would require less effort than processing a passage with underlined key words.

The relational adjuncts were also of interest for exploring the kind of processing that fosters inferential learning. To support inferential learning, perhaps it is sufficient to identify for the student the fundamental and relevant relations embodying the new scientific concepts. If this were so, then the experimenter-provided labeled schematics (if not the student-generated schematics), by highlighting and elaborating the relevant relations, should be sufficient to enhance inferential learning (e.g., see Mayer & Gallini, 1990). An
extension of this position based on MAP is that enhanced inferential learning would be obtained only for the literal texts; for the analogy-based texts, the schematic adjuncs would produce processing already invited by the text itself and therefore would not improve inference performance.

Method

Participants. A total of 201 undergraduates attending Purdue University were tested in partial fulfillment of a course requirement. Twenty-five participants were assigned randomly to each of the eight conditions formed by the factorial combination of text type (literal and analogy) and study adjunct (control, key-word highlighting, student-generated schematic, and experimenter-provided schematic), with the exception of the analogy control condition to which 26 participants were assigned.

Materials. For each of 12 scientific concepts (gleaned from various high school and college texts), we used short didactic texts developed by Donnelly and McDaniel (1993). There were two basic texts for each concept: one provided a literal description of the concept and the other, in addition, provided an analogical statement relating the concept to a domain more familiar to learners. In the key-word conditions, we identified words or phrases that referred to key aspects of the concept. These key words were the same for the analogy and literal text material, and they were highlighted by both underscoring and boldfaced type. For the experimenter-provided schematic conditions, we developed labeled schematics that pictorially represented the relational dynamics of the conceptual information in the text. The same schematic was used for the analogy and literal texts. In the student-generated schematic conditions, students were provided with adequate drawing materials (paper and drawing pencils).

The multiple-choice test questions used to assess factual and inferential learning were also taken directly from Donnelly and McDaniel (1993). There were four multiple-choice questions associated with each of the 12 scientific concepts (literal and analogical). In all cases, the first two questions were basic level. These questions merely required the student to recognize factual information that was read in the concept. The last two questions were inference level. These questions were designed so that one could reason from the familiar domain in the analogy to answer the question correctly. For example, in the concept "collapsing star," the familiar domain was ice skaters who pirouette faster as they object's speed would decrease, which was the correct inference to answer the question that they would be answering. The concept of an analogy was described to participants in the analogy condition. It was noted that analogies map information from a source domain onto a target domain. Pilot data revealed that such explicit instructions greatly facilitate participants' understanding and use of analogical passages, a finding that is common in this literature (see Bisanz, Bisanz, & Lefevre, 1984; Donnelly & McDaniel, 1993; Gentner, 1989; Gick & Holyoak, 1980; Reed, Ernst, & Banerji, 1974; Vosniadou & Schommer, 1988).

All participants were informed that they would receive 12 short texts, in sets of 6, and that each set was to be followed by a series of multiple-choice questions. The order of set was counterbalanced across participants. After participants studied the first set of six concepts in self-paced fashion (study time was not recorded because of the size of the testing-session groups), a short distractor task involving the reading of a short story and rating 63 statements on a scale of 1-5 for level of importance was assigned. After reading the second group of concepts, students were distracted with 32 questions from Paivio and Harshman's (1983) verbalizer-visualizer imagery preference questionnaire. The distractor tasks were included to avoid possible ceiling effects in performance.

Following each distractor task, participants answered 24 alternative multiple-choice questions regarding the six preceding scientific concepts (4 questions per concept). Half of these questions were factual-level questions that required strict retrieval of the stimulus materials, and half of the questions were inference-level questions that required the participant to speculate or go beyond the information given. The questions for each concept were grouped together with the 2 factual questions presented first followed by the 2 inference questions. Participants were given 10 min to answer these questions. Participants then proceeded to the second group of concepts, distractor task, and multiple-choice questions.

Because one participant in the analogy key-word highlighting condition did not follow instructions by neglecting to provide an answer for all of the multiple-choice test questions, he or she was replaced with a participant from the same pool. All other participants provided an answer for every test question, as instructed. Nevertheless, to ensure that replacing the one participant did not unduly change the pattern of results, we recomputed the analyses reported in the Results section. In these analyses we excluded the lowest scoring participant in each of the other seven experimental conditions. The pattern of the means and the effects that emerged as significant completely paralleled those reported in the results.

In developing the analogies, one of the authors designed several possibilities for each target concept. From these possibilities, an honors student in physics was asked to select the best, the next best, and the worst analogy for each concept and in addition was asked to evaluate the appropriateness of the analogy selected as best. This student believed that all of the "best" analogies appropriately captured the critical relations of the target concepts. Even though it remains possible that in a few instances alternative analogies might have "worked" better than the ones developed for the study, the set of analogies used in this study appears to be reasonable and representative of appropriate analogies.

It is perhaps possible that this procedure could have influenced the general performance levels for inference-type (or factual) questions; however, as detailed in Footnote 4, the particular type of example questions used had no apparent effect on test performance.
Results

The proportions of correct responses for factual-level and for inference-level questions were tabulated for each participant. The means are displayed in Table 1. Donnelly and McDaniel (1993, Experiment 1) established that participants (N = 120) sampled from the same population used in the current experiment produced substantially higher performance (at least 50% higher) than the Donnelly and McDaniel no-exposure controls, suggesting that the instructional treatments generally produced learning. The theoretically motivated predictions developed in the introduction focused on comparisons of text type and study adjuncts for each question type. Accordingly, to provide the most straightforward and coherent test of these predictions, we analyzed the data from each question type with separate (by dependent measure) two-variable, between-subjects analyses of variance (ANOVA). In these ANOVAs, text format (analogy and literal) and encoding condition (control, key-word, subject-generated schematic, and experimenter-provided schematic) were the variables. The familywise alpha was set at .05 for each source of variance.

Factual-level questions. There was no difference in correct responding in general as a function of text format (F < 1) or as a function of encoding condition (F < 1). These patterns were qualified, however, by a significant interaction between text format and encoding condition, F(3, 193) = 3.22, MSE = 0.016. The nature of this interaction is revealed in comparisons associated with the theoretical predictions outlined in the introduction. Specifically, for a particular text format, we compared each control condition with each of the encoding-adjunct conditions using Dunnett’s multiple comparison procedure (using one-tailed Dunnett values because of the directional nature of the predictions). First, as anticipated and as shown in Table 1, adding encoding adjuncts to the literal condition produced no improvement in encoding of factual-level information. In fact, the literal text alone produced factual learning that was nominally better than with any adjunct (none of these comparisons were statistically significant).

Second, as predicted, key-word highlighting significantly improved factual-level performance for the analogical text, whereas adding schematic encoding to the analogy condition did not produce significant improvement. A possible caveat in interpreting the obtained advantage of the key-word highlighting for the analogical text is that it may reflect a units-of-analysis (e.g., Levin, 1992) artifact of the testing arrangement. Specifically, because all students in a particular encoding adjunct condition were tested in just two experimental sessions, the just-mentioned result may reflect an idiosyncratic testing session for the key-word group. This seems unlikely because (a) there was no parallel key-word advantage for the literal text (a condition that was tested in the same sessions as the analogical text) and (b) all participants were tested in the same room, by the same experimenter, and at approximately the same time of day (early evening). Nevertheless, we calculated means for each particular testing session for the analogy text (thereby obtaining two means for each encoding-adjunct condition) and found that performance across sessions for any particular encoding adjunct was quite stable (each pair of means differed by no more than .015). Most important, both means in the key-word highlighting condition were greater than any of the six means from the other three conditions. The probability of this occurring by chance is less than .05, using an exact permutation calculation (p = .036, one-tailed). Thus, this more conservative analysis supports the conclusion that adding key-word highlighting significantly improved factual-level performance for analogy text.

We performed one additional comparison with just the control conditions (conditions with no encoding adjuncts) to allow direct correspondence with previous work. The literal text did not produce statistically more accurate responding than the analogical text (.82 vs. .77). It is worth noting, however, that the magnitude and direction of the nominal difference is not out of line with a previous report showing a small but reliable advantage of the literal text over the analogical text for factual-level information (in Donnelly & McDaniel, 1993, literal favored analogy with differences ranging from .04 to .06 across four experiments).

Inference-level questions. There was a significant main effect of text format, F(1, 193) = 4.22, MSE = 0.026, with analogical text producing higher performance than literal text (.69 vs .64). This effect did not interact with encoding adjunct (F < 1), nor did the encoding adjunct variable significantly influence performance (F < 1).

Table 1

<table>
<thead>
<tr>
<th>Dependent measure/</th>
<th>Control</th>
<th>Highlighted key word</th>
<th>Generate schema</th>
<th>Provided schema</th>
</tr>
</thead>
<tbody>
<tr>
<td>text type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literal</td>
<td>.82</td>
<td>.76</td>
<td>.81</td>
<td>.79</td>
</tr>
<tr>
<td>Analogy</td>
<td>.77</td>
<td>.85</td>
<td>.78</td>
<td>.81</td>
</tr>
<tr>
<td>Inference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literal</td>
<td>.67</td>
<td>.65</td>
<td>.60</td>
<td>.66</td>
</tr>
<tr>
<td>Analogy</td>
<td>.69</td>
<td>.68</td>
<td>.69</td>
<td>.71</td>
</tr>
</tbody>
</table>

Discussion

A main finding was that an analogical text can be supplemented with an appropriate study adjunct (in this case highlighting key words) to increase the level of fact learning produced by analogy-enhanced scientific text. The significant increase in fact learning due to highlighting key words was accompanied by intact inference learning in the analogical-text condition. In contrast, for the literal text, not one of the three study adjuncts examined here (either experimenter or student generated) was able to boost the inference learning for the new scientific constructs. Thus, gains in inference learning appear difficult to come by, so
far being limited to analogy-enhanced text, at least with these materials (see Donnelly & McDaniel, 1993, for similar findings with additional adjuncts).

From a theoretical standpoint, these data are consistent with the general thrust of the MAP analysis of learning strategy effectiveness (McDaniel & Einstein, 1989). Specifically, not just any study adjunct will successfully combine with analogy-enhanced text to enhance fact learning. As anticipated by the MAP analysis, only the key-word highlighting condition, a condition that focused learners on the factual details, statistically improved factual learning when combined with the analogy-based text. The study adjuncts that presumably encouraged learners to attend to interrelations among the elements comprising the scientific constructs did not statistically enhance factual learning. According to the MAP analysis, this is because such relational processing was already invited by the analogy embedded in the text. Similarly, the key-word highlighting did not enhance factual learning of the literal text because for such texts learners are processing the details in the hope of extracting or understanding whatever important relations they can (cf. Einstein et al., 1990; Graf & Levy, 1984).

The results are also pertinent to other issues. We thought that pictorial schematics might increase inference learning based on previous research and on the theoretical notion that inference learning might be supported by more complete processing of the interrelations of the elements of scientific constructs. The present findings reinforce a cautionary note in terms of embracing schematics as a general way to promote inference learning, though schematics might be effective in some bounded cases. Mayer and Gallini (1990) showed that labeled schematics enhanced learners' performance on problem-solving questions (questions requiring inferencing and extension of the conceptual material in the text), provided that the schematics explicitly labeled the component parts and the behavior of each part. In line with Mayer and Gallini, our schematics were labeled as completely as possible and each text was presented adjacent to the schematic. However, our texts in their entirety were shorter than the verbal material (labeling) associated with the effective schematics in Mayer and Gallini (see Figure 1 in their article), and Mayer and Gallini used a longer text (750 words). It may be then that schematics are more effective if applied to more detailed and elaborated constructs (longer texts). Perhaps the types of inference questions are critical as well; we used multiple-choice questions, whereas Mayer and Gallini used short-answer questions. It is clear that more work is needed to track down this issue.

From an applied perspective, this experiment demonstrates that increases in comprehensive learning (both factual and inference learning) might require a rather specific combination of study adjunct and text: combination of the highlighted key-word adjunct and the use of analogy. Perhaps of more central importance for educational application, the gains from this combination were quite modest. The highlighted key-word condition produced only nominally better factual learning than a literal condition alone, and the gains in inference learning from the analogy conditions, though statistically significant, were not startling. This pattern raises the question of whether there remains some study adjunct or combination of analogy and a study adjunct that would produce more robust gains in either factual learning or inference learning or both.

Experiment 2

A study adjunct that has been generally successful in improving learning is the elaborative interrogation method. The elaborative interrogation method is a higher order questioning strategy (e.g., relative to other kinds of typically used adjunct questions that require learners to only report facts from the text; see Pressley, Tanenbaum, McDaniel, & Wood, 1990) that requires learners to explain why phenomena described in the text occur (Willoughby, Wood, & Khan, 1994). In earlier work with this technique, students were asked to write a simple factual statement and to answer a why question afterward to clarify the relation between the subject and predicate (Pressley, McDaniel, Turnure, Wood, & Ahmad, 1987). The elaborative interrogation students showed substantial increases over reading control groups in memory for the facts, and this finding has been replicated for a wide range of students (Seifert, 1993).

Further relevant to the present purposes, some research has been supportive of a knowledge-base explanation of the benefits of elaborative interrogation (Willoughby, Waller, Wood, & MacKinnon, 1991). This explanation suggests that elaborative interrogation produces superior factual learning because it activates relevant prior knowledge and fosters connections between this prior knowledge and the new facts. The activation of prior knowledge to help understand new relations might also aid in inference learning. Moreover, the particular knowledge presumably activated through elaborative interrogation is idiosyncratic for each learner, thereby perhaps allowing it to become more useful to the learner (in contrast to analogy, for which all learners might not have sufficient prior knowledge to extract and reason with the fundamental relations expressed in every analogy; cf. Pressley et al., 1987, for similar reasoning with regard to limited effects of experimenter-provided elaborations). Accordingly, we thought it possible that elaborative interrogation would prove to be a potent learning adjunct for fostering both factual and inference learning. In testing this idea, we addressed the following three fundamental issues.

First, though the use of questioning techniques is increasingly advocated in educational settings (Hansen, 1994; King, 1994), many experimental studies have not examined the benefits of elaborative interrogation with materials that would necessarily be representative of educational contexts. For instance, participants are typically presented with lists of facts to learn, with these facts presented as sets of sentences (e.g., Pressley et al., 1987; Pressley, Symons, McDaniel, Snyder, & Turnure, 1988) or arranged into paragraphs (Woloshyn, Willoughby, Wood, & Pressley, 1990). To our knowledge, only one published study (Seifert, 1993) has examined the learning effects of elaborative interrogation with more naturalistic prose passages (i.e., passages
containing main ideas, detail facts, intersentential relations, etc.). Testing sixth- and seventh-grade participants, Seifert found enhanced learning for the main ideas of the text, when those main ideas were targeted by why questions for which the participants had to generate an answer. Elaborative interrogation did not enhance memory for supporting details. Thus, there is modest, though sparse, evidence that elaborative interrogation can improve factual learning from text, at least for children. The following experiment builds on Seifert’s initial study by examining the use of elaborative interrogation for adult factual learning from technically difficult text.

Second, Seifert (1993) also assessed inferencing performance and, using a one-item test, found no better performance for the elaborative interrogation condition (relative to an underline control condition). This failure to find improved inference learning due to elaborative interrogation must be considered preliminary, however, because as Seifert noted, a one-item test is a weak assessment. Moreover, even if the finding were to hold for children, there is as yet no evidence on whether inference learning would be improved for adults who engage in elaborative interrogation while processing technical prose. This experiment provides the first test of this important issue.

A third objective was to explore some possible explanations of the benefits of elaborative interrogation, if found in the present context. As mentioned earlier, one favored account is that elaborative interrogation activates prior knowledge that can be connected to the target facts so as to increase their memorability (Martin & Pressley, 1991). This explanation does not necessarily hold in all situations, however, because elaborative interrogation improves learning of facts for which students have no prior knowledge (Woloshyn, Pressley, & Schneider, 1992). To gauge the extent to which elaborative interrogation stimulated activation of prior knowledge for the present materials, we evaluated students’ responses to the why questions.

If there were little or no evidence of prior-knowledge activation in the presence of learning benefits (due to elaborative interrogation), then what might explain these benefits? Few alternatives have been clearly specified (Woloshyn et al., 1992). With text materials like those used here, perhaps elaborative interrogation forces learners to extract a more coherent text-based representation, which would support better memory performance (Kintsch & van Dijk, 1978). Because recall is assumed to be sensitive to the coherence of text representations (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983), in this experiment we added a recall test to our learning assessment procedure. If an elaborative interrogation condition produced higher recall levels, then it is consistent with the idea that with text materials, elaborative interrogation produces a more coherent or integrated representation. If recall were not enhanced (by elaborative interrogation) and if activation of prior knowledge were not evident (in learners’ elaborations), then other explanations of learning effects of elaborative interrogation would bear consideration.

There was one other feature of this experiment that needs mention. We were concerned that our short texts might not provide enough background context to support effective elaborative interrogation, so we added some brief general background material to each text to facilitate answering why questions. We presented these modified texts both with and without elaborative interrogation so that any effects of the added material could be assessed independently from the effects of the elaborative interrogation.

**Method**

**Design and participants.** The design was a $2 \times 3$ between-subject factor with text condition as one independent variable (literal or analogy-based text) and instructional adjunct as the other variable. The three levels of the instructional adjunct variable were control (no addition to the basic texts), modified (background material added to the basic texts), and elaborative interrogation (why questions added to the modified texts). The Appendix gives an example of the two sets of materials (for “Collapsing Stars” and for “Earth’s Rotation”) representing the factorial combination of the two variables. Twenty-eight participants were randomly assigned to each of the two control conditions and the literal text with elaborative-interrogation condition; 29 participants were assigned to the analogy-based text with elaborative interrogation condition; and 30 participants were assigned to each of the two modified-text conditions. Participants were from the same population as in the previous experiment.

**Materials.** The 12 basic texts were the same as those used in Experiment 1. The modified texts contained an additional sentence or phrase that provided additional background for answering the why questions. In constructing the why questions, we took care to avoid duplicating the multiple-choice test questions. The questions were also designed to encourage learners to go beyond explicitly stated information to provide answers but did not necessarily require learners to do so (the Appendix gives examples of the why questions).

**Procedure.** The procedure for the conditions without the elaborative interrogation adjunct was identical to that used in Experiment 1. For the elaborative interrogation conditions, each why question was printed immediately below the appropriate text, and blank space was provided for participants to write in their response to the question. As in Experiment 1, participants proceeded through the texts at their own pace, after which they answered factual and inference questions about the material conveyed in the texts. After reading and answering the multiple-choice questions to all 12 of the texts, participants were given a surprise recall test. To signal the contents to be recalled, we provided participants with the first few words of each text. The order in which the texts were cued corresponded to the original presentation order. Finally, participants were given a brief questionnaire asking about their scholastic training in mathematics and science. Participants responded with (a) the number of courses in the physical sciences taken in college; (b) the number of those courses taken in high school; (c) the number of mathematics courses taken in college; and (d) the number of mathematics courses taken in high school.

---

4 Inadvertently during the instructions, approximately half of the participants in the literal control group received a factual question for their example instead of an inference question. This had no impact on test performance, as performance was equivalent for participants in this group regardless of the kind of example presented (Fs < 1 for the comparison of participants receiving a factual example question with those receiving an inference example question on both factual and inference test questions).
Results

The proportions of correct responses to factual and inference level questions were analyzed with separate 2 (literal text and analogical text) × 3 (control, modified, and elaborative interrogation) between-subjects ANOVAs. These are reported first, followed by the recall results and the evaluation of the elaborations generated in the elaborative interrogation condition (see Table 2 for means). The familywise alpha was set at .05 for each source of variance.

Factual-level questions. Literal texts produced significantly better factual learning than analogical texts, F(1, 167) = 9.14, MSE = 0.02. The main effect of instructional adjunct was also significant, F(1, 167) = 13.87. Examination of Table 2 reveals that the instructional-condition effect was entirely due to an increase in factual learning for elaborative interrogation compared with the control and the modified conditions; these latter two conditions produced identical performance. Finally, the enhancing effects of elaborative interrogation held regardless of the kind of text (literal or analogy; F < 1).

To allay the concern that the benefit of elaborative interrogation may have been due to idiosyncratic effects of one of the test sessions (cf. Experiment 1 results), we calculated means representing each testing session for each of the instructional conditions (thereby yielding two means per condition, when collapsed across text type). There was little variation across testing sessions (in the elaborative interrogation condition when collapsed across text type). There was little variation across the pairs of means representing each testing session for each of the instructional conditions; these latter two conditions produced identical performance. Finally, the enhancing effects of elaborative interrogation held regardless of the kind of text (literal or analogy; F < 1).

Table 2

<table>
<thead>
<tr>
<th>Dependent measure/ text type</th>
<th>Instructional adjunct</th>
<th>Modified + elaborative interrogation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factual</td>
<td>Control</td>
<td>Modified</td>
</tr>
<tr>
<td>Literal</td>
<td>.75</td>
<td>.74</td>
</tr>
<tr>
<td>Analogy</td>
<td>.70</td>
<td>.69</td>
</tr>
<tr>
<td>Mean</td>
<td>.72</td>
<td>.72</td>
</tr>
<tr>
<td>Inference</td>
<td>Control</td>
<td>Modified</td>
</tr>
<tr>
<td>Literal</td>
<td>.58</td>
<td>.60</td>
</tr>
<tr>
<td>Analogy</td>
<td>.60</td>
<td>.58</td>
</tr>
<tr>
<td>Mean</td>
<td>.59</td>
<td>.59</td>
</tr>
<tr>
<td>Recall*</td>
<td>Control</td>
<td>Modified</td>
</tr>
<tr>
<td>Literal</td>
<td>.51</td>
<td>.37</td>
</tr>
<tr>
<td>Analogy</td>
<td>.47</td>
<td>.39</td>
</tr>
<tr>
<td>Mean</td>
<td>.49</td>
<td>.38</td>
</tr>
<tr>
<td>Recall of analogy</td>
<td>Control</td>
<td>Modified</td>
</tr>
<tr>
<td>Analogy</td>
<td>.52</td>
<td>.41</td>
</tr>
</tbody>
</table>

*Recall proportions reflect recall of information common to all text conditions.

As above, means for each test session were calculated, and as before there was little variation across the pairs of means (greatest difference was .028). Each of two means in the elaborative interrogation condition (collapsed across text type) was greater than any of the other means (p = .067, one-tailed, by an exact permutation calculation). An ANOVA (not collapsing over text type) treating each mean as an individual subject confirmed the significant instructional-condition effect, F(2, 6) = 24.11, p < .005, MSE = 6.28, and the absence of any other effects.

Recall. For each of the 12 concepts, the idea units common to the literal and the analogical texts were identified. The modified and elaborative interrogation groups contained an extra idea unit (or two) relative to the control group. Therefore, to keep all groups comparable, we scored only those idea units that were common to the three different conditions. For the majority of concepts, this meant that one extra idea unit was omitted. Two judges who were unaware of the purposes of the study independently scored the recall protocols for the presence of these idea units. The same judges scored the protocols for all of the conditions; across conditions, the judges agreed in their ratings on 90% of all idea units recalled. Discrepancies were resolved by a third judge (also unaware of the conditions).

These data were submitted to a 2 (basic text and analogical text) × 3 (control, modified, and elaborative interrogation) between-subjects ANOVA. The main effect of instructional adjunct was significant, F(2, 167) = 5.82, MSE = 0.03. This effect was due to significantly lower recall in the modified condition than the control and elaborative interrogation conditions, which themselves did not differ (as shown by Newman–Keuls tests). More important, for present purposes a planned test showed that recall in the elaborative interrogation condition was not significantly higher compared with the two control conditions combined, F(1, 167) = 1.58. There were no other significant effects (Fs < 1).

Next, we analyzed recall of the idea units pertaining only to the analogy (in the analogy conditions). This scoring was
done by the same two judges who scored the earlier data, and there was 99% agreement by these judges in scoring the analogy idea units. A one-way ANOVA again revealed a significant effect of instructional condition, \(F(2, 84) = 3.11, \text{MSE} = 0.09\). However, in this case elaborative interrogation was associated with the poorest recall. This decrement in recall was statistically significant when compared to the control condition but not when compared to the modified condition (as shown by Newman–Keuls tests). Thus, the recall patterns are not supportive of the idea that elaborative interrogation produced better factual and inference learning due to extraction of a more coherent text representation.

**Answers to why questions.** The content of responses to the why questions was analyzed. We classified the responses for each answer given by each participant into one of two categories, text-based or prior knowledge (correctness of the response was not taken into account). Response content that was judged to be text-based merely restated or paraphrased information that was given in the target concept. Response content that included any additional information not present in the target concept was judged to be based on prior knowledge. A judge unaware of the purpose of the experiment scored the responses according to this classification scheme. Nine individual responses appeared ambiguous to the judge, and adjudication of these responses was determined jointly by us and the judge.

The results showed clearly that participants’ responses were almost entirely constructed from the text itself. On average, participants’ answers reflected prior knowledge on less than 1 out of the 12 passages read (\(M = 0.63\)), and this varied minimally as a function of whether analogy was included (\(M = 0.70\)) or not (\(M = 0.57\)). To gain a more detailed picture, we examined the use of prior knowledge for the participants with the most extensive science backgrounds. Participants who reported completing at least 10 science courses in high school and college were identified as “experts.” Ten participants met this criterion, and for these 10, evidence of prior knowledge increased (in their answers), but not dramatically so: Fewer than 2 out of 12 answers (\(M = 1.5\)) included new information (not available from the text). The mean for the remaining 47 participants was 0.45.

**Discussion**

The results of this experiment extend previous research on the elaborative interrogation technique in a number of ways. First, performance on the factual questions showed that requiring adult learners to answer why questions about technical prose can effectively increase the level of factual learning from that prose. The few other studies that have investigated elaborative interrogation for adult learners studying prose materials have reported similar benefits (e.g., Woloshyn et al., 1990). However, the prose materials used in those studies were basically lists of facts arranged into paragraphs (see Seifert, 1993). In the present experiment the texts, though brief, were arguably more representative of “real world” (instructional) prose.

Additionally, the benefit of elaborative interrogation for factual learning was widespread, extending to information not necessarily targeted by the why questions. In previous studies, the why questions have been targeted specifically at the to-be-learned facts. Seifert (1993) in addition tested for learning of facts in the text not targeted by the why questions and found no increased learning of those facts in relation to a no-elaborative interrogation control. Seifert, however, sampled younger learners, whereas adult learners were sampled in the current experiment. Also, Seifert used longer texts than used herein. The conditions under which why questions can enhance learning of nontarget information warrants further study. It is clear, however, that enhanced learning of information not directly targeted by the why question can be added to the growing list of learning benefits produced by elaborative interrogation.

Second, and of central interest to this study, why questions significantly increased inference learning. The one previous experiment that examined inference performance after elaborative interrogation failed to find such a benefit (Seifert, 1993), but as noted earlier that experiment probably did not adequately test inference-level performance. The enhanced inference learning due to elaborative interrogation is especially impressive in light of the failure in Experiment 1 to find benefits for other reasonable instructional adjuncts and the limited benefits from analogy (see also Donnelly & McDaniel, 1993). Moreover, the increases in inference learning were just as evident for literal as for analogy-based texts.

Perhaps even more interesting, the benefits of elaborative interrogation were apparently obtained without extensive activation of prior knowledge (as indicated by the content of participants’ responses to the why questions) or without more extensive recall of the information contained in the texts. Indeed, with elaborative interrogation, participants recalled substantially less of the analogical information than those participants in the control-analogy condition. More important, this decrement in recall of analogical information was not offset by superior recall of the other content (see Table 2). Thus, it appears that elaborative interrogation did not simply produce storage of more textual content per se or produce a more integrated representation of the text (in terms of relating the propositions together in a coherent text base). If a more integrated text base had been produced, recall would presumably have benefited (Kintsch & van Dijk, 1978). In the next section we offer a tentative explanation for the learning increases produced by elaborative interrogation, and we consider the implications of this study for characterizing the components of learning that lead to appropriate inferencing with newly acquired scientific concepts.

**General Discussion**

The present experiments examined the effectiveness of a variety of techniques—analogy, analogy with key-word highlighting, labeled pictorial schematics, and elaborative interrogation—for enhancing inferential thinking with
newly acquired scientific concepts. As outlined previously, there are theoretical bases for expecting that all of these techniques might promote learning that supports inferencing with newly learned concepts, at least to some extent. Up to this point, however, these techniques have not been extensively investigated in this regard, nor have they ever been directly contrasted in one study so as to gauge their relative effectiveness.

The results were clear. Labeled schematics, whether learner-generated or experimenter-provided, did not enhance learning, either in terms of inference-level performance or in terms of factual learning. Analogy produced small increases in inference learning in comparison with a literal presentation (significantly so in Experiment 1, but not Experiment 2), but with a concomitantly small cost in terms of factual learning (significant in Experiment 2, but not Experiment 1). As anticipated by our theoretical analysis, factual learning with analogical texts tended to improve when these texts were supplemented by highlighting key words. In contrast to the other techniques, elaborative interrogation produced substantial learning benefits both for inference-level and factual-level performance.

From an applied standpoint, this pattern implies that comprehensive learning of scientific concepts (factual as well as inference learning) presented in texts like those used here can best be facilitated (considering just the adjuncts investigated here) through an elaborative interrogation study technique. Also, the pattern suggests caution for instructional design in terms of combining treatments thought to enhance learning in the hopes of gaining even more. In particular, analogy did not enhance inference-level performance in relation to that obtained with a literal text when learners were required to engage in elaborative interrogation. It is clear that the present texts were much less involved and elaborate than those used in educational situations, and it may be that analogy or labeled schematics would be more effective with more extensive "lessons." On the other hand, the present texts were more educationally representative than the sets of individual facts typically used in the laboratory to demonstrate the potential educational benefits of elaborative interrogation (Woloshyn et al., 1992). Thus, this study provides some of the strongest evidence yet reported that the elaborative interrogation study technique merits serious consideration for use in academic contexts. Further specification of the situations in which elaborative interrogation, as well as analogy, might be most beneficial rests on consideration of the factors that underlie the effects of analogy and elaborative interrogation. Accordingly, we turn now to possible explanations for the learning outcomes associated with analogy and elaborative interrogation.

Several explanations have been proposed for the mnemonic benefits (increased factual learning) of elaborative interrogation, including activation of congruent prior knowledge, heightened arousal, increased attention to content, and the process of generating information associated with the why question (i.e., a generation effect) or induction of processing that is well matched to the criterial task (see Pressley et al., 1987, and Woloshyn et al., 1992, for amplification of these ideas). The idea that the positive effects of elaborative interrogation represent a generation effect (Pressley et al., 1987) does not appear to be supported. If requiring learners to generate information per se enhanced performance, then generating labeled schematics (in Experiment 1) should have produced improvements over a literal-text control. This result did not occur, consistent with previous work showing that generation of elaborations per se did not produce better factual learning over a reading control (Martin & Pressley, 1991).

Regarding the heightened arousal view, a priori there is no reason to expect that answering why questions would be more arousing than generating labeled schematics (the other condition that required active construction of additional information). Nor is it likely that increased attention per se to content accounts for the benefits of elaborative interrogation in the present circumstances. Increased attention to at least some of the content of the instructional texts was presumably induced by adding labeled schematics and by underlining key concepts in the control literal text (Experiment 1), yet these conditions produced no improvement in either factual or inference learning relative to the literal control (see Table 1). More directly, elaborative interrogation did not enhance recall and, in some cases, attenuated recall of the content. Enhanced recall might be expected on an account that appeals only to increased attention. A prior-knowledge explanation also does not appear able to account for the finding of improved learning subsequent to answering why questions. We found little evidence of use of prior knowledge in participants' answers to the why questions. Given the apparent shortcomings of existing explanations, we offer a tentative account based on current theoretical treatments of text comprehension. These theories suggest that one outcome of understanding textually presented information is the construction of a situation model, a mental representation of the situation described by the text (rather than a representation of the text base itself; van Dijk & Kintsch, 1983). Another way to describe this is that the reader constructs a mental model of the "world" described by the text. Such a mental model would allow in essence a mental simulation of the scientific constructs, one end product of which would be the ability to infer beyond the stated facts (Johnson-Laird, 1983). Thus, it might be that the cognitive challenges presented by elaborative interrogation (in terms of reasoning with the information presented in the text) force the learner to construct a more complete situation model than he or she would otherwise construct (see King, 1994, p. 17). In light of the recall results in Experiment 2, it is important to note that richer situation models do not necessarily produce higher recall levels for the text's contents (Perrig & Kintsch, 1985).

The foregoing analysis leaves unanswered, however, why embedding analogy in instructional text would fail to improve inference performance to the extent that elaborative interrogation does. There are at least two possible reasons for the relative ineffectiveness of analogy in the present study. Because the analogies were provided by an external agent, they either (a) might have not always been consistent enough with the learner's knowledge base or (b) were not
rich enough to help the learner construct a better understanding (mental model) of the to-be-learned scientific concepts (Zook & Maier, 1994).

The present study was not explicitly designed to illuminate the factors underlying the effects of the selected study adjuncts. The findings do suggest possible interpretations, but most important, they emphasize, first, the potentially limited utility of analogy for promoting learning from text that conveys new scientific concepts, even when analogy is supplemented with additional study adjuncts. This conclusion converges with recent work indicating that there are a host of learner and instructional factors that lead to analogical misconceptions (e.g., Zook & Maier’s, 1994, model encompasses six factors), thereby disrupting potential benefits of analogy. A second central finding, in stark contrast to the limited effects of analogy (and labeled schematics), is the potency of elaborative interrogation for stimulating learning that includes both acquisition of factual information and acquisition of information that supports inferencing beyond the stated facts.

References


Donnelly, C. M. (1995). *Crucial cognitive and personality factors in the creative process*. Unpublished manuscript, Purdue University, West Lafayette, IN.


Appendix

Example Materials and Test Questions

The why question at the end of each passage appeared only in the elaborative interrogation condition (Experiment 2). The under-scoring and boldface type within each passage appeared only in the highlighted key-word condition (Experiment 1). In addition, the passages with the why questions are the modified passages used in Experiment 2. The passages with the schematics are those used in Experiment 1. Adapted from “Use of Analogy in Learning Scientific Concepts” by C. M. Donnelly and M. A. McDaniel, 1993, Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, p. 987. Copyright 1993 by the American Psychological Association.

Literal

**Collapsing Stars**

A body in motion has constant energy and energy must be conserved. It takes less energy to spin as an object’s size gets smaller. Collapsing stars spin faster and faster as they fold in on themselves and their size decreases. This phenomenon of spinning faster as the star’s size shrinks occurs because of a principle called “conservation of angular momentum.”

Why does an object speed up as its radius gets smaller (as in conservation of angular momentum)?

Analogy

**Collapsing Stars**

A body in motion has constant energy and energy must be conserved. It takes less energy to spin as an object’s size gets smaller. Collapsing stars spin faster as their size shrinks. Stars are thus like ice skaters who pirouette faster as they pull in their arms. Both stars and skaters operate by a principle called “conservation of angular momentum.”

Why does an object speed up as its radius gets smaller (as in conservation of angular momentum)?

[See Figure A1 for the schematic used in the experimenter-provided schematic condition.]

**Collapsing Stars**

1. Which of the following is true about collapsing stars?
   a. The smaller they are, the slower their rotation.
   b. The smaller they are, the slower their orbit.
   c. The smaller they are, the faster their rotation.
   d. The smaller they are, the faster their orbit.

2. Which of the following terms applies to the way collapsing stars rotate?
   a. The principle of conservation of angular momentum.
   b. The principle of conservation of static equilibrium.
   c. The principle of conservation of centripetal force.
   d. The principle of stellar parallax.

3. What would happen if a star “expanded” instead of collapsing?
   a. Its rate of rotation would increase.
   b. Its rate of rotation would decrease.
   c. Its orbital speed would increase.
   d. Its orbital speed would decrease.

4. What would happen if a collapsing star first expanded then contracted?
   a. It would first increase its orbital speed, then decrease orbital speed.

[See Figure A1 for the schematic used in the experimenter-provided schematic condition.]

**Figure A1.** Collapsing stars spin faster as their size shrinks. Stars are thus like ice skaters who pirouette faster as they pull in their arms. Both stars and skaters operate by a principle called “conservation of angular momentum.”
b. It would first decrease its orbital speed, then increase orbital speed.
c. It would first increase its speed of rotation, then decrease speed of rotation.
d. It would first decrease its speed of rotation, then increase speed of rotation.

**Literal**

**Earth's Rotation**

The Earth remains in a constant state of rotation. The spinning Earth rotates in the solar system at a constant tilt. This tilted spinning, which is called *precession,* is due to a balance between the spin of the Earth which acts to *straighten* the Earth up, and gravity which acts to *pull* the Earth down.

Why is the Earth at a tilt?

**Analogy**

**Earth's Rotation**

The Earth remains in a constant state of rotation. The spinning Earth behaves like a slowing toy top that wobbles at a tilt (but never topples). This tilted spinning, called *precession,* is due to a balance between the spin which *straightens* Earth (or the top) up, and gravity which *pulls* Earth (or the top) down.

Why is the Earth at a tilt?

[See Figure A2 for the schematic used in the provided schematic condition.]

**Earth's Rotation**

1. The phenomenon called *precession* occurs when an object such as the Earth,
   a. spins on its axis.
   b. orbits another object, such as the Sun.
   c. spins on its axis at a tilt.
   d. orbits another object, such as the Sun, at a tilt.
2. Why does the spinning Earth *precess* in space?
   a. Because spinning *straightens* it up, but gravity *pulls* it down.
   b. Because spinning *pulls* it down, but gravity *straightens* it up.
   c. Because spinning *pulls* it toward the Sun, but gravity *pushes* it away from the Sun.
   d. Because spinning *pushes* it away from the Sun, but gravity *pulls* it toward the Sun.
3. What would happen if the Earth stopped spinning?
   a. The Earth would fall on its side.
   b. The Earth would *straighten* up.
   c. The Earth would freeze at its original tilt.
   d. The Earth would tilt at a different angle.
4. What would happen if there were no gravity?
   a. The Earth would continue to rotate, which would cause it to fall on its side.
   b. The Earth would continue to rotate, which would cause it to *straighten* up.
   c. The Earth would stop rotating and so it would fall on its side.
   d. The Earth would stop rotating and so it would freeze at its original tilt.