Abstract

Ninety-six undergraduates at the University of Virginia studied a 2866 word excerpt from a tenth grade biology textbook either in a standard typographical layout or in a special one containing typographical and spatial cues. Both groups studied the text for about nineteen minutes. (Results of paraphrastic recall, a multiple choice test, and a questionnaire suggested that the typographical and spatial cues facilitated learning and memory.) Potential classroom applications were discussed.

(The format used here is similar but not identical to the one we used.) Here we used bold face letters for emphasis; in the experiment we used capital letters because it was easier on a typewriter. Here the authors determined the segmentation units; in the experiment groups of college students determined them.

The instructions for using our format are: (The passage is printed in a special format using line spaces, brackets, and bold face letters. Line spaces set off each idea, and brackets set off the gist of each idea, whether or not the idea is important. When an idea is important, its gist is in bold print.)

When we read a text to learn, we assimilate its semantic core, often remembering it in our own words instead of in its original surface representation. [Whether we realize it or not, however, surface representation influences our reading and our memory for what we read.] For example, readers find answers to true-false questions faster when a text provides spatial cues about its grammatical phrase structure (Frase & Schwartz, 1979; Hartley, 1980) and sentences are reread faster when the first and second presentation have the same typography (Kolars, 1975). The latter result suggests that surface representation is an integral part of memory and sets the stage for developing typographies that facilitate memory.

[In the present study, we endeavored to develop typographical and spatial cues that would improve learning and memory of a 2866 word excerpt from a tenth grade biology textbook.]
Background

[The foundation for our study was laid by a two-year project, "Learning from Math and Science Textbooks," conducted at the University of Virginia by (in alphabetical order) James Deese, Thomas Estes, John Rotondo, Wayne Shebilske, and M. Elizabeth Wetmore.] The aim of our project has been to provide a theoretical and empirical foundation for the development of better textbooks and better diagnostic and instructional techniques for teaching the kinds of reading demanded by textbooks in mathematics and science.

[The first two years were spent studying interrelationships between subject factors, formal text structures, perceived text structures, comprehension strategies, and the understanding of conceptual content in typical classroom reading assignments.]

[We will briefly review our large on-going project here, because it determined the specific typographical manipulation used in the present study.]

Our research is based on the assumption that what a reader sees in a text is as important as what an author writes in it. A major thrust in the early phase, therefore, was the development of methods for measuring perceived text structure. [Two procedures especially important to the present study were dividing the text into idea units and rating the importance of those units.]

Measuring Perceived Idea Units. We had one group of subjects divide our biology text according to the following instructions:

Now that you have read this passage, we would like for you to go back through it and divide it into what we call "idea units." [An idea unit encompasses a complete thought; and while idea units often coincide with sentences or are set off by punctuation, this is not always the case. More than one idea unit may occur in a single sentence; one idea unit may carry over from one sentence or paragraph to the next.]

Sometimes you may not be certain where one idea ends and another begins. We would like for you to make judgements despite this uncertainty. Please turn back to the passage that you just read and put slash marks (/) after each idea unit. Raise your hand if you have questions.

Figure 1 shows some of our units. [They tend to be considerably larger than those obtained by other researchers who asked subjects to divide a passage where ever it was natural to pause during reading] (Johnson, 1970; Brown & Smiley, 1977; Frase and Schwartz, 1979).

If you try to mark pausal units in Figure 1, you will see that our units often contain more than one pausal unit, and our idea units often terminate at the end of a pausal unit. You should also be able to mark larger units in Figure 1 by combining some idea
Typographical and spatial cues

units. For example, the first four and the last two might form two groups corresponding to examples and main points. [In fact, some subjects who divided our passage according to the above instructions marked smaller or larger units than those shown in Figure 1.]

[Despite variability in coarseness, however, idea units tended to be congruent (i.e., smaller units from fine grain analyses tended to be nested within larger units from coarser analyses) suggesting that smaller units may be used as building blocks for larger units when subjects structure an internal representation of text.]

We consider the units shown in Figure 1 to represent only one level of that whole hierarchical structure.

Importance Ratings: 1= unimportant
10= important

TWO METHODS OF CLASSIFICATION

(The basic idea of classification is not difficult to understand. We all do some informal classifying,) and almost anything may be classified--stamps, rocks, clouds, even the kinds of weather.

(The words in a dictionary are classified. They are classified according to their spelling—that is, alphabetically).

(In classifying objects we could use an alphabetical method, arranging them according to the alphabetical order of their names.) Suppose a supermarket manager arranged his merchandise alphabetically. Think of the varied goods to be found under the letter A: abalones, almonds, apples, apricots, artichokes, and many more. These would be followed by bacon, baking powder, beans, beef, beets, bread. Imagine the practical difficulties in such a system. Refrigerators for perishable groceries would have to be scattered throughout the store.

(Actually, in any supermarket we find that the merchandise has generally been grouped according to the nature of the product). In one section we find various kinds of canned goods; in another, fresh fruits and vegetables, in a third, meats. Moreover, each of these sections may be further divided. Familiarity with this system of classification enables the shopper to locate groceries easily and quickly.

(Thus WE CAN CLASSIFY IN EITHER OF TWO WAYS: ACCORDING TO LIKENESSES IN NAMES OR ACCORDING TO LIKENESSES IN OBJECTS THEMSELVES.)

(FOR BIOLOGICAL CLASSIFICATION NAMES OF ORGANISMS ARE CERTAINLY OF MUCH LESS IMPORTANCE THAN CHARACTERISTICS); so the alphabetical method is not satisfactory.

Figure 1. An excerpt from Biological Sciences and Curriculum Study, High School Biology (Green Version), 2nd ed., Chicago: Rand-McNally, 1963. The excerpt is printed with the following typographical and spatial cues: (1) Idea units are separated by one line space; (2) phrases and sentences corresponding to the gist of idea units are set off by parentheses, and (3) important gist statements are capitalized. Our modified format maintained paragraph indentations that were in the original. In the experiment the passage was typed on 8.5 x 11 in. pages in pica with blank 1 inch margins on all sides. Here the right margin is used to show importance ratings.
[Taking advantage of between-subjects variability in coarseness of unitizing, Rotondo (Note 1) developed a single-linkage clustering algorithm for determining the perceived hierarchical structure of a text.] In principle, we thought that it would be best to use the entire hierarchical structure in any analysis of a text.

In practice, however, we needed to choose one level of analysis to make our initial work more manageable. Rotondo therefore proposed a rationale for choosing one level of a text’s hierarchical structure. [He computed the mean number of idea units used by subjects and then chose the hierarchical level that came closest to having that number of units, resulting in the units shown in Figure 1.]

[We then had a second group of subjects rate the importance of those units.]

*Measuring Perceived Importance of Textual Units.* Figure 1 shows some importance ratings provided by college students who were given the following instructions:

The author of the passage you just read had main ideas that he or she wished to convey. Some of the ideas in the passage are very important with respect to the author’s main points, while others are much less important. [We would like you to tell us how important you perceive each idea to be with respect to the author’s main points.]

On the following pages the ideas of the passage are listed with two rating scales next to each. The first scale is for you to say whether the idea is *important unimportant* to the author’s main points. Even if you aren’t sure of the choice you are making, rate each idea as either *important* or *unimportant*.

A rating of *A* will mean you think the idea is *important*.

A rating of *B* will mean you think the idea is *unimportant*.

The second scale is for you to say how *sure* you are of your choice on the first scale.

*A* will mean you are *highly sure* of your choice between important and unimportant for this idea.

*B* will mean you are *sure* of your choice between important and unimportant for this idea.

*C* will mean you are only *slightly sure* of your choice between important and unimportant.

*D* will mean you are *unsure* of your choice between important and unimportant.

*E* will mean you are *highly unsure* of your choice between important and unimportant.
By comparing ratings of college students with those of tenth graders, we demonstrated an advantage of measuring perceived text structure over the more common practice of limiting textual analysis to formal structures (e.g., Frederiksen, 1975a,b; Norman & Rumelhart, 1975; Anderson & Bower, 1973; Kintsch, 1974; Schank, 1973; Grimes, 1972). [Tenth grade importance ratings were highly correlated with college ratings, but they deviated systematically, in that tenth graders overestimated the importance of concrete examples and underestimated the importance of abstract main points.]

[High school teachers who inspected our results agreed with the college students’ assessment of importance, and they were surprised by the tenth grade results.] The teachers had known that tenth graders have trouble determining the main points in a text, but they had not foreseen the existence of systematic differences between themselves and their students. We had not anticipated a systematic difference either, but we were enthusiastic about it, believing that we were on the right track.

Measuring Learning and Memory. [We were encouraged further when we used our normative idea units and importance ratings to analyze learning and memory.]

[Teachers often given short answer tests to determine whether or not a student understood the vocabulary, facts, main ideas, and inferences related to a text.]

[Many teachers have recognized the potential advantage of adding an open-ended question such as “Recall what you have read in your own words,” but they have avoided them because of the difficulty in evaluating the answers.]

[We found that we could make good use of behaviorally defined units and importance ratings in evaluating answers to open-ended questions.]

A group of college students and a group of tenth graders read our tenth grade text in preparation for a difficult examination including both essay and objective tests. After a single study session, students paraphrased the text including as much detail as possible. In one of several recall analyses, we determined whether or not students included the gist of each idea unit. The proportion of college students recalling an idea unit was significantly correlated with importance ratings ($r = .39, p < .001$). Tenth grade recall was correlated with college recall ($r = .68, p < .001$), but it also deviated systematically. [Specifically, tenth graders tended to recall relatively unimportant concrete examples at the expense of important abstract main points, corroborating the pattern observed in importance ratings.]

Our finding of significant correlations between importance ratings and recall agrees with results of Johnson (1970) and Brown and Smiley (1978). [Instead of emphasizing the positive correlation itself as they did, however, we laid stress on the large amount of recall variance unaccounted for by importance ratings.] The correlations for college students accounted for only 15 percent of the variance and the correlation for tenth graders was worse. We concluded that both groups have difficulty in identifying and
recalling important ideas and we decided to do something about it by teaching the students and by improving the textbooks.

[The present study is part of our effort to enhance the communicative effectiveness of textbooks.]

Rationale

After reviewing conflicting reports on the influence of a text’s layout on comprehension, Hartley (1980) called for “comprehensive studies comparing different kinds of reading tasks . . . and different layouts for different kinds of text” (p. 76). Our project provides a framework for such a comprehensive approach, but first [we need answers to the following preliminary questions: Can special typographical designs adequately portray the structures we had measured? If so, will those typographies improve learning and memory? Finally, will gains in learning and memory manifest themselves in paraphrastic recall?]

[Our first attempt at developing a special format based upon measures of perceived text structure consisted of the following manipulations: 1) Idea units were separated by one line space; 2) phrases and sentences corresponding to the gist of each idea unit were set off by brackets, and 3) important gist statements were capitalized.] Figure 1 shows the resulting format for our tenth grade biology passage.

[We compared reading rates and comprehension measures of college students who read this passage in either a standard layout or in our special layout.]

Methods

Subjects. [Subjects were 96 undergraduates at the University of Virginia who participated for pay or for partial fulfillment of a course requirement.]

Design. [We used a simple between-subjects design comparing performance of 47 students who read the standard format with 49 students who read the special format.]

[We measured reading time, paraphrastic recalls and scores on a detailed multiple choice test. We also asked students to evaluate our typographical design.]

Procedure. We collected data in a large auditorium. For practical reasons, we ran the standard group first. [Both groups read instructions for studying, read the material, wrote paraphrases of the text, took a multiple choice test, and then answered a questionnaire.]
[In order to test the effectiveness of the special format itself we minimized differences in instructions between the two groups.] The instructions for studying are shown below:

This booklet contains an excerpt from a biology textbook. Please work under the assumption that you are reading this material for a class in which you will be given a detailed examination. Specifically, the exam will include 34 multiple choice questions as well as essay questions on the material. With that assumption in mind, please study this excerpt as you would normally study such material during your first study session with it. *The passage is printed in a special format using line spaces, brackets, and capital letters. Line spaces set off each idea, and brackets set off the gist of each idea, whether or not the idea is important. When an idea is important, its gist is capitalized,* (*-* omitted for standard layout group). You may do anything you normally do when you study such as underline or take notes. Of course, you will not be able to look at your notes or the text during the test.

Please proceed at whatever rate you would ordinarily use during your first study session with this kind of material. Since people vary a great deal in their reading rates, and since different booklets contain different passages, you need not worry if others seem to be going faster or more slowly than you.

When you finish studying the passage, record your time, and raise your hand.

If you have any questions about these instructions, please raise your hand now. If not, turn the page, record your time, and begin studying.

All other instructions and procedures were identical for the two groups.

[A large digital clock which counted from 0 to 9999 seconds was used to time each part of the experiment.] Subjects wrote the four digits displayed when they started and when they finished each part. Because of a clock malfunction, we got times for only 15 of the students in the standard layout group.

Results

[It is important to keep in mind that the results reflect the state of comprehension after an initial study session.] Most students study a homework assignment once and then return to it one or more times to review before an exam (Shebilske, Fisher & Karmiohl, Note 2). During an initial study session, students may go over material several times, but the initial session is nevertheless only part of the student’s total study effort. In effect we have taken a “snapshot” of the first stage of the total study process.

[As a result scores are lower than one would expect to see for a class who had completed the entire study process.]
The results suggest that students who read the special layout learned and remembered the text better without spending appreciably more time studying. Average reading times were 18 min. 29 sec., and 19 min. 5 sec. for the standard and special layouts respectively ($t(62)=0.89$, $p>.05$). Average proportions of idea units recalled were .17 and .23 for the standard and special layouts ($t(94)=2.42$, $p<.01$). We will report more detailed analyses of the paraphrastic recall followed by analyses of the multiple choice test and the questionnaire.

Relationship Between Importance and Paraphrastic Recall. The correlation between importance ratings and recall increased from $r=.39$ for the standard layout to $r=.50$ for the special layout, thus the percentage of variance accounted for by importance increased from 15% to 25%. Even though this increase is not statistically significant ($z = .83$, $p>.05$), it suggested that the special layout may have brought recall into better linear alignment with importance by improving recall to a greater extent for important material.

We analyzed this possibility by testing recall separately for unimportant (ratings less than 6) and important (ratings greater than 6) idea units. We found that the proportion of unimportant ideas increased insignificantly from .12 for the standard layout to .15 for the special layout ($t(94)=1.49$, $p>.05$), while the proportion of important ideas recalled increased significantly from .22 for the standard layout to .29 for the special layout ($t(94)=2.50$, $p<.01$).

A further breakdown of the analysis into ten separate importance categories showed that the special layout group did better in every category and significantly better only for idea units rated 5 (.07 vs. .11; $t(94)=3.40$, $p<.001$), 8 (.23 vs. .37; $t(94)=3.40$, $p<.001$) and 9 (.47 vs. .59; $t(94)=1.99$, $p<.05$).

[Apparently, the typographical and spatial cues, improved learning and memory of important information without reducing the learning and memory of unimportant information.] The same conclusion is suggested by the multiple choice test.

Multiple Choice Results. Scores on the 29 question multiple choice test were highly variable, ranging from 28% to 93% within the standard layout group and from 41% to 81% within the special layout group. Consequently, [no differences between groups were statistically significant. The results are interesting however, because they follow the same pattern as the paraphrastic recall.] The average score for the standard format was 59% in comparison to 63% for the special group. An item analysis revealed that of the 21 questions that tested important ideas, 15 yielded higher scores for the special format group and 6 went in the other direction. Of the 8 questions that tested unimportant ideas, four yielded higher scores for the special format group.
Typographical and spatial cues

Again, the low scores reflect the fact that students only completed an initial study session with the material. [A comparable group of students who were allowed to take the same passage home to study and review, spent a total of 67 min. with the passage on the average, and they obtained a mean score of 75% on the same test.]

[As most students know, they are not well prepared for a test after an initial study session. The important point here, however, is that they seem to be better prepared when the text provides spatial and typographical cues.] The survey of attitudes towards the special format supports this conclusion.

Student Evaluations of the Special Format. Our subjects answered the questions shown in Table I and then explained their answers. Six students were eliminated because their explanations indicated that they understood “special format” to mean the whole experimental situation. Six other failed to answer the questions. Of the 36 remaining, [a clear majority thought that the special format helped both during reading and during recall.]

Table I
Responses to a questionnaire on whether the special format seemed to help or to interfere during reading and during recall.

1. Circle the number on the scale below that best indicates how the special format influenced your ability to comprehend the text during reading.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Interfered a great deal</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Interfered</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Interfered slightly</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Neither helped nor interfered</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Helped slightly</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>Helped</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>Helped a great deal</td>
<td>1</td>
</tr>
</tbody>
</table>

2. Circle the number on the scale below that best indicates how the special format influenced your ability to recall what you read in the text.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Interfered a great deal</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Interfered</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Interfered moderately</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Neither helped nor interfered</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>Helped slightly</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>Helped</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>Helped a great deal</td>
<td>6</td>
</tr>
</tbody>
</table>
Most of the comments were very favorable and very telling. [The 30 out of 36 students who favored the special format during reading agreed that the typographical and spatial cues helped them allocate their processing resources more effectively.] Some comments were as follows:

"It helped to read the text faster. I was not worried about skipping important points. It was helpful in reviewing."

"After the first time I read and studied it I could later use the format to remember the more important points. Because the major points followed a logical order so the next could be easily memorized."

"It showed a difference between thoughts and prepared me for this change in thought. I categorized the ideas which helped slightly."

"It helped in the understanding by: organizing the thoughts picking out the important parts and helping me in visualize what the process was in the reading."

The 33 students who favored the special format during recall made comments such as the following:

"I could see the pages in my head and I remembered the capitalized letters during the test."

"The capital letters stood out in my mind."

"Major ideas stood out in my mind in almost an outline form. Things were easy to recall."

"After the important points were learned, the gist of the ideas would come naturally into mind."

"The format seemed very organized. The emphasized points were recalled faster than the small detailed points."

"It again allowed my mind to concentrate only on the main and important points. It stopped the cluttering of my memory."

"Especially the words that were capitalized stood out in my memory of the text and increased my recall ability."

"I am sure that I remembered more than I would have in a first study had it not been presented this way."

[Some of the comments suggest that the typographical and spatial cues were an integral part of memory and that subjects could improve their recall by]
visualizing the text. Other comments leave open the possibility that recall was better simply because the typographical and spatial cues helped students to study more efficiently.]

Either way, [both objective results and subjective evaluations suggest that our special format helps students learn and remember text.]

The six who were against our layout made comments such as "I like to pick out what I believe is important and not be forced to see what someone else thinks is important. Part of the learning process is figuring out what is important." [It would have been interesting to analyze data separately for those who opposed the format, but there were too few of them to make the analysis worthwhile.]

Our results suggest, however, that analyses of individual differences will be an essential part of future research.

Discussion

Learning from textbooks is troublesome even for college students who are intelligent enough to master a text’s content and literate enough to recognize the words and sentences (Shebilske, 1980). [We therefore hope to develop more effective textbooks and better instructional techniques for teaching the special study skills demanded.] The present experiment convinces us that typographical and spatial cues can play an important role in that endeavor.

[Our project is not sufficiently developed to support specific applications, but during this embryonic stage we must keep in mind the possibility of such applications.] With respect to printing costs, our special layout would be more practical than those requiring that texts be completely rearranged (e.g., Wendt, 1979). Measuring perceived text structure would be expensive however, and it is hard to imagine publishing houses shouldering this expense for college textbooks.¹

¹The expense of measuring perceived structure could be reduced by using small panels of expert readers instead of large groups of student readers. The authors’ segmentation of the present text is a step in that direction. We used intuitions that we developed by observing many student segmentations of other texts. We have not yet validated our intuitions by comparing our units with those generated by a group of naive students. We plan to do so with other passages.

This passage does not lend itself to such a comparison. It discusses various possible levels of segmentations, relationships between levels, and functional significance of levels. Asking students to segment this text might be like asking students to count how many times they think about pink elephants in the next hour. Surely this passage has perceptual units and students occasionally think about pink elephants. But how does one obtain a valid measure of either? This limitation is minor since few passages are about segmentation.
Perhaps the best application, therefore, would be in teaching younger students. [Tenth grade science teachers might use our typographical design, along with special lesson plans, to teach students how to zero in on important points.]

If this skill could then be transferred, students could learn to cope with upper level books printed in standard formats.

[Before any practical recommendations can be justified, however, we must increase the effectiveness of our format.] We will conclude this paper with three plans for doing so.

*Training Students to Use Typographical and Spatial Cues.* Other experiments on special typographical designs have provided short lessons prior to the experiment (e.g., Wendt, 1979). We did not do so in the present study because we wanted to avoid the various control groups that would have been required to ensure that improvements were caused by the typographical design instead of by the lessons. Furthermore, we did not know what to teach. What is the best way to use our typographical and spatial cues? Our questionnaire indicated that individual students adopted different approaches. For example, some read all the capitalized words first and then went back to fill in details; others read the whole passage first and then reviewed only the important ideas. [With larger samples, we hope to identify the most effective strategies. We will then be prepared to develop lessons, which will hopefully increase the effectiveness of our typographical format.]

*Portraying Perceived Hierarchies.* [Another plan is to represent more levels of the perceived hierarchical structure.] Following the lead of Hartley and Burnhill (1976), we will vary the amount of vertical space between adjacent idea units in proportion to the frequency with which the boundary between the units is chosen as a division point by subjects segmenting the text; the larger the frequency, the greater the space separating the units. We did not represent more levels in the present study because we think this more subtle spatial cue will require prior training, which we wished to avoid for reasons already stated. In addition, we have only recently developed a procedure for measuring perceived hierarchies for individual students. Thus, we will soon learn more than we presently know about the advisability of portraying normative hierarchical structures.

*Pursuing a Comprehensive Approach.* [Our textbook project at the University of Virginia provides a comprehensive framework for evaluating typographical designs.] We are studying a variety of textbooks and various task demands so that we will be prepared to analyze typographical designs with respect to a wide range of conditions.

[One variable that seems especially important is discourse style, which determines, among other things, the location of critical information.] Deese (1980) points out that while any classification of text is an oversimplification, scientific texts can be classified as one of two kinds: deductively structured and inductively structured. Deductive structures state general principles at or near the beginning of a section and then explain or exemplify that principle. Inductive structures present examples first
and then bring the reader to a general principle, which is usually stated near the end. The tenth grade passage used in the present study tends to be inductive, which is the predominate style in lower level textbooks. Inductive style can be seen in Figure 1 where our special typography shows at a glance that the main points are located near the end.

[Because our typographical design manifests style, we think it will be especially effective when used in conjunction with lessons on how to read different styles.]

Such lessons are in the future. For now, however, [we have taken a first step toward integrating typographical design into our total effort to improve textbooks and to teach the kinds of study skills demanded by textbooks.]

Reference Notes


References

Frederiksen, C. H. Effects of context induced processing operations on semantic information acquired from discourse. *Cognitive Psychology*, 1975b, 7, 139-166.


**Acknowledgments**

The work reported here was supported under Grant 00233 under the joint sponsorship of NSF and NIE. We thank James Deese, Thomas Estes, and M. Elizabeth Wetmore for helpful comments on the study reported here. We also thank Adèle Blennerhassett, Evangeline Motley, and Paul Twohig for help in data collection and analysis.