ABSTRACT

The amount and structure of prose learning derived from vocabulary definition and graphic post-organizer construction methods of independent study tasks is examined. Undergraduate subjects (n=144) read a 2000-word passage in a 20-minute reading period. Two subject groups performed text information tasks while a third acted as a control to measure recall derived from a reading-only strategy. Two days later, all subjects, including an additional control group (n=44) who had neither read nor studied the material, completed a multiple-choice test on the passage. Results of a multivariate analysis of variance indicated that both study procedures were successful, but that the two groups differed on specific types of information learned. A single discriminant function was found to separate the two study groups. Vocabulary rehearsal led to the learning of greater amounts of subordinate level information, while graphic post-organizer usage led to the learning of more superordinate information. This study was limited due to low internal consistencies, short test drawn from a single content area, and learning measured on the basis of recognition only. (Author/PN)
Specific Learning Outcomes Attributable to Study Procedures

Timothy Shanahan
University of Illinois at Chicago
College of Education - Box 4348
Chicago, Illinois 60680

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
T. Shanahan

TO THE EDUCATIONAL-RESOURCES INFORMATION CENTER (ERIC)."

Running Head: Rehearsal strategies
ABSTRACT

Specific Learning Outcomes Attributable to Study Procedures

The purpose of this study was to examine the differential impact of two study techniques upon type of information learned. Subjects (n=144) were randomly assigned to four groups. Subjects read a 2000-word passage. Group I studied definitions of 24 technical terms from the text. Group II studied the interrelationships of these same terms using a graphic post-organizer technique. Group III was a control group that only read the passage, but did not rehearse any of the information from the text. Group IV was also a control group; they neither read the passage, nor rehearsed the information. Two days later, subjects completed a specially designed multiple-choice test. Results of a multivariate analysis of variance indicated that both study procedures were successful, but that the two groups differed on specific types of information learned. A single discriminant function was found to separate the two study groups. Vocabulary rehearsal led to the learning of greater amounts of subordinate level information, while graphic post-organizer usage led to the learning of more superordinate information.
"The conventional measure of sheer amount of overall information recalled is inappropriate for research on learning from text, since the type of information that is learned ... can vary under different treatments without changing the overall recall level" (Dee-Lucas & DiVesta, 1980). The purpose of this study is to examine and compare both the amount and types, or structures, of prose learning which can be derived from two methods of study (vocabulary definition and graphic post-organizer construction). This investigation attempts to define the structure or type of learning in a manner that is both logically consistent with the learning tasks, and theoretically relevant to models of text and text processing (Kintsch & van Dijk, 1980; Meyer, 1975). This study is unique in that it examines the influence of independent study tasks, as opposed to teaching procedures, upon the types of learning derivable from natural text. It examines types of information and training tasks not previously considered in research of specific learning outcomes.

Task Appropriate Learning

Rehearsal strategies refer to those tasks that require readers to do more thorough processing of text information than would result from the use of a reading-only strategy. Tasks such as underlining, rereading, precis writing, and repetition are examples of rehearsal strategies. These strategies have been proposed and investigated because of their potential usefulness as teaching and study tools. However, studies of the effectiveness of many of these strategies have provided inconsistent results (Anderson & Glover, 1981; Annis & Davis, 1975; Arnold, 1942; Brady & Rickards, 1979; Davis & Annis, 1978-79; Fowler & Barker, 1974; Idstein & Jenkins, 1972; Noall, 1962; Orlando & Hayward, 1978; Rickards & August, 1975; Riley & Dyer, 1979; Summers, Forester & Jeroski, 1978).
A major reason that rehearsal strategies have not consistently been found to result in increased learning and retention might be the use of different and, in some respects, unequal dependent measures across studies. That is, these studies have differed in the amount of disparity which has existed between the rehearsal tasks and the dependent measures. Thus, subjects have appeared to learn more when the dependent measures have closely resembled the study tasks than they have when there has been a greater discrepancy between the training tasks and the outcome measures. (For two theoretical perspectives on why such mismatches would lead to discrepant outcomes see discussions of encoding specificity (Tulving & Thomson, 1973) and schema theory (Rumelhart & Ortony, 1977)).

Much of the evidence, which suggests that "different forms of activity result in the construction of somewhat different representations of reality" (Olson, 1977, p. 69) is, unfortunately, indirect. Such evidence suggests the possibility of differences without demonstrating a direct link between the processing activities and the learning outcomes. Studies of learning outcomes attributable to selective information cueing provide a good example of this type of research (Frase, 1972; Gagné, 1978; Rickards, 1979; Rothkopf, 1972). When, prior to reading, students are given specific questions about the text, it has often been found that they are, after reading, able to answer similar questions quite well. However, these same students find it difficult to answer other text questions, about non-cued information. It is usually concluded that subjects, in such experiments, narrow the focus of their processing, and that they encode only that information that appears to be within the focus of the original questions. Similarly, subjects do better on memory tests when their intention to learn is manipulated by informing them of the nature of the test (i.e., recall, recogni-
tion) to be taken (Tversky, 1973; Tversky, 1974). Such research evidence sug-
gests that processing shapes not only the amount of learning, but also the type
of information that is learned. Nevertheless, it infers the nature of process-
ing on the basis of outcome. More compelling evidence would be obtained if pro-
cessing differences were manipulated and, as a result, learning differences
occurred.

Direct studies of structural differences in learning outcomes have rarely
examined text learning. Instead, these studies have focussed on list learning,
picture recall, mathematical computation, problem-solving, etc. (Bransford,
Nitsch & Franks, 1977; Maier, Thurber & Janzen, 1968; Mayer & Greeno, 1972;
Nitsch, 1977). In this research, groups of subjects have been trained on a
variety of tasks. Learning has then been measured on multiple tests, essen-
tially identical to each of the training tasks, either in form or content. These
studies have differentiated the structure of learning on the basis of how the
information was to be used, or how it was to be learned (i.e., recall, recogni-
tion, problem-solving, computation). It has usually been found that different
types of information, as categorized by usage, are recalled as a result of dif-
ferent training activities.

During the past decade, the nature of information in text, especially its
hierarchical organization and its cohesiveness, has received much atten-
Although these analyses have increased our understanding of text structure, they
have not often been used to distinguish the types of learning which result from
different teaching and study strategies. Only one study of specific learning
Outcomes were found which distinguished types or levels of text information remembered as a result of using different study or training procedures (Dee-Lucas & DiVesta, 1980): Subjects in this study were asked to read a specially constructed text. They were then directed to rehearse certain types of text information by one of three procedures. One group of subjects attempted to write down all of the information they could recall ("recall of facts"). The second group attempted to match a list of attributes with their paragraphs of origin ("recognition of specific facts"). The third group attempted to fit ideas from the passage into the top levels of a prefabricated hierarchical tree ("recall of superordinate information"). Subjects then were examined on multiple tests, each test being identical to the rehearsal procedures of one of the three groups. This study found that subjects learned different types of information (i.e., detail vs. superordinate) dependent upon the type of study method used. Subjects did best when they were tested in the same way that they had rehearsed. There were also interactions between the various study and test types indicating that tasks can direct subjects' attention to types of information, independent of how the information is to be used (i.e., recall, recognition).

The Dee-Lucas & DiVesta study indicates that the same usage classifications identified in previous learning experiments, could be employed for describing text learning. This study also found other, text specific, information distinctions that could be useful for differentiating task specific learning outcomes. This study was not without its limitations, however. The text that was read by subjects in this study was atypical. It was constructed specifically to enable the subjects to study it in the ways specified. It is possible that these
information distinctions would not be as useful for describing the learning which would result from the use of similar learning strategies with other text materials. Secondly, the classifications of superordinate information and factual or detail information does not represent a clear distinction. As operationalized in this study, a detail could be selected from either high (a superordinate position) or low (a subordinate position) in the information hierarchy. Superordinate information could only be drawn from the upper ranks of the hierarchy. Finally, although the subjects in this study read and rehearsed the text information independently, it should be noted that, except for group one, their rehearsal was dependent upon the availability of previously constructed study materials. Thus, the learning tasks were not methods that subjects could readily use for studying.

**Graphic Post-Organizers**

What is needed are investigations of task specific learning outcomes which focus on the use of independent study strategies with natural text. One such strategy which might profitably be examined with a specific learning outcome paradigm is graphic post-organizer construction. A graphic post-organizer (GPO) is a visual arrangement, or tree diagram, of the technical vocabulary of a text. The GPO represents the semantic relationships which exist between the vocabulary terms in the text. The construction of GPO's has been found to be a particularly powerful rehearsal strategy (Barron & Stone, 1974; Barron, 1980; Barron & Scwhartz, 1979). GPO construction requires students, after reading a text, to attempt to reconstruct, or to generate, the hierarchical and parallel relationships described or implied in the text.
The use of GPO construction as a rehearsal strategy is particularly interesting both because of its proven effectiveness and its theoretical relevance. Theories of learning and memory storage often suggest that learners must construct something like an internalized GPO in order to remember information (Ausubel, 1963; Rumelhart & Ortony, 1977). It is quite possible that the GPO strategy helps learners to build up such an internalized schematic representation of the text information, in a manner that allows the information to be recalled easily. It also seems possible, however, that GPO construction might require the rehearsal of certain types of information, to the detriment of the learning of other types of information. Studies of GPO construction have usually evaluated the effectiveness of this strategy using tests designed to assess knowledge of relationships only. It is possible that this strategy enhances "relationship" learning at the expense of other, more micro-propositional, forms of learning. Also, because of its emphasis upon the hierarchical organization of information, it might encourage the recall of superordinate information to the detriment of the recall of subordinate information. This study is designed to evaluate these possibilities.

Method

Subjects

The participants in this study were 72 undergraduates currently enrolled in a teacher preparation program at a large midwestern university. These subjects were assigned to one of three study groups, and they were each paid $5.00 for two days participation. An additional 44 subjects took the test only, and they were paid $1.00 each. These additional subjects were randomly drawn from the
same subjects pool as those subjects assigned to the study groups. The test-only subjects took part in the experiment for one day, only.

**Materials and Tests**

Subjects were asked to read and study a 2000-word chapter on cell theory. This passage was selected because: (a) It used a large number of technical vocabulary terms. (b) It contained material judged to be relatively unfamiliar to most of the participating subjects. (c) It was written at a low level of difficulty (i.e., 7th grade level) as measured by the Harris-Jacobson readability formula (1981). This suggests that the word recognition demands of the passage were such that they should have provided little or no barrier to recall or understanding.

A multiple-choice test consisting of 32 items designed to measure recall of information from the cell theory chapter was administered to all subjects. This instrument was designed so as to ask equal numbers of questions about information at different levels of the information hierarchy (i.e., high or superordinate, low or subordinate). The investigator and a trained assistant independently created graphic post-organizers for this chapter, using all of the technical vocabulary from the passage. Vocabulary terms which were included in the top 40% of both post-organizers were considered to be high in the information hierarchy. These terms were then used to design high information level questions. Similarly, terms included in the bottom 40% of both of the graphic post-organizers were judged to be low in the information hierarchy, and these were used to design low information level questions. Questions were also included to assess recall of vocabulary and relationship information. Vocabu-
lary questions asked about single vocabulary terms (i.e., definitions). Relationship questions required that students relate two or more vocabulary terms (i.e., categorizations). The test questions were presented in random order, as selected from the following categories:

- 8 items—high level information—relationships
- 8 items—"" level information—vocabulary
- 8 items—low level information—relationships
- 8 items—"" level information—vocabulary

Two items from each category represented material from the chapter which was read but not studied. These items were included as a check of whether the rehearsal techniques had an impact on recall. (Data from one of the high level-relationship questions was lost due to an inadvertent clerical error. For this reason, all comparisons across categories of information used proportions of answers correct and not raw scores).

This test had a reliability (KR-20) of .75. Of course, the specific information test sections used in this analysis, because of their shorter lengths, had lower reliabilities (ranging from .55 to .65). Given the low internal consistencies of these tests, effect sizes attributed to the various study outcomes represent conservative estimates of these effects.

Procedures

On day one of the experiment, one of three study packets was randomly assigned to each subject. All subjects first read the cell theory chapter during a 20-minute reading period. Subjects were to read the chapter as if they were to be examined on it for one of their courses. If they finished reading
before the end of the allotted time, they were to indicate that they had finished reading the passage and they were to reread the passage until time was up.

All subjects read the chapter in its entirety within the time limits.

**Group I:** Subjects in this group (n=26) were told to write definitions from memory for each vocabulary term in a list of 24 provided by the investigator. If subjects were unable to generate a definition from memory they were to look back at the text.

**Group II:** Subjects in this group (n=25) were given the same 24 vocabulary terms, in the same order, though written on 3x5 cards and not in list form. They were to organize these cards to represent the hierarchical and parallel relationships between them. An example was given. Again, if students were unable to place all of the terms into the hierarchy, they were to look back at the text as necessary.

**Group III:** The subjects in this group (n=21) were a control group. These subjects were to complete a page of 3-digit multiplication problems. This group was included so as to allow the measurement of recall derived from a reading-only strategy.

These rehearsal tasks were completed during a 12-minute study period. Another period, also 12-minutes in duration, was provided. During this time, subjects were to combine into dyads, within groups. Subjects were told to compare and contrast the results of their efforts during the study period. Although students were informed that they were not required to achieve consensus during these discussions they were to attempt to understand the source of any
disagreements. Study packets were then collected and subjects were told not to discuss the study techniques, nor the content of the text material.

Two days later, all subjects, including an additional control group (n=44) made up of subjects who had neither read nor studied the material, completed the test on cell theory. This group allowed a base-line measure of knowledge of cell theory to be made for subjects drawn from this population.

**Hypotheses**

This study was designed to test the following hypotheses:

1. Subjects who studied the passage, by either method, were expected to perform significantly better on a test of cell theory knowledge than would the reading-only subjects or the test-only subjects. Similarly, the reading-only subjects were expected to outperform those that had not read the passage. These hypotheses were intended to test the effectiveness of the study techniques over a reading-only strategy.

2. Subjects who studied the material, by either technique, were expected to do significantly better on the material that they had studied than on the material that they had read only. Again, this hypothesis was intended to test the efficacy of the study procedures in raising the amount of text learning.

3. The two study groups were expected to perform equally well on overall text score, as both rehearsed the information for the same amount of time and their performance was being measured on a test which assessed learning of different types of information to an equivalent degree.
4. The study groups were expected to be significantly different in their performances on the various subtests (relationship questions; vocabulary questions; high level questions; low level questions), as a result of the different study procedures used.

**Results**

Mean proportions and standard deviations for performance on the various question sets for each group are reported in Table I.

The purpose of this study was to compare the types of information learned as a result of the two rehearsal techniques. However, before examining this it was necessary to be certain that the study activity actually enhanced learning. An analysis of variance was performed to compare overall performance on studied information for three groups (the two study groups were pooled for this comparison). The results of this analysis indicated that the three groups differed significantly in test performance ($F(2, 114) = 25.74$, $p < .01$). A post hoc comparison of means, using the Scheffé procedure was then made. This analysis indicated that the studiers performed significantly better than did those subjects who only read the passage; and subjects who read did better than the control group subjects who only were tested ($p < .05$).

Another way of considering the effectiveness of the two study procedures is to compare performance on those items that were studied with performance on those items that were not studied. On average the students in the study groups answered about 13% more of the studied questions than they did the
non-studied ones. These scores were compared using a one-tailed single sample t-test \((t = 5.34, df = 50, p < .01)\). The rehearsal techniques significantly enhanced learning.

Use of the rehearsal techniques resulted in learning, but there was no difference between the two study groups on overall test scores \((t = 1.48, df = 49, p > .30)\). This lack of difference was expected, as the test used to measure learning was equally balanced between the various information types. If subjects from either group did especially well on any parts of the test, it was predicted that lowered performance on the other sections would counterbalance such results on the overall test score.

In order to find out whether these groups differed on any of the categories of information, a 4 (dependent measures: scores on the high, low, relationship, and vocabulary questions) X 1 (independent measure: group membership) multivariate analysis of variance (MANOVA) was performed. This analysis is based on the assumption that the dispersion (variance-covariance) matrix for each group is similar. To insure equality of dispersion matrices, Box's M (Amick & Crittendon, 1975) was calculated \((M = 41.75, F (30, 22706.7) = 1.29, p > .10)\). Given that the null hypothesis, that group dispersions do not differ, was not rejected, it was assumed that these groups had similar dispersion matrices. Thus, any differences found between groups in the MANOVA would be due to the power of the dependent measures to discriminate among group centroids.

In calculating the MANOVA it was discovered that so much collinearity existed between the dependent measures that the error matrix would not
invert. This is a common occurrence with ipsative measures. For this reason, only the high information, low information, and vocabulary information items were used as dependent measures to test the hypothesis of equality of centroids. Measures were selected for inclusion in this analysis in order of the magnitudes of their univariate F-ratios. (Univariate F's are reported in Table II. Although none of them were statistically significant, it should be noted that the intercorrelations among the dependent measures were relatively low, and thus, these could combine to explain significant variance in group membership)

The results of this analysis indicate that the two study groups differed in their performance on the dependent measures. The group centroids were different. \( F(3, 47) = 2.82, p < .05 \). The vocabulary group and the GPO group were distinguishable on the basis of the types of information that were learned from the rehearsal techniques.

Finally, a two-group stepwise discriminant function analysis was performed, with minimum F-ratio required for any variable to enter the analysis set at 1.0 to select only those variables that optimally discriminated between groups. Two of the variables (high level and low level information) made significant contributions to the differentiation among groups. The two variables together yielded a Wilks' Lambda of .8754, which is approximated by a multivariate F-ratio of 3.41 \( (p < .05, df = 2, 48) \). This indicates that the GPO and vocabulary groups are distinguishable on the basis of their performance on questions designed to tap different levels of the information hierarchy. The GPO group did slightly better on the high level
or superordinate information questions, while the vocabulary rehearsal group
was superior on the low level information. (Neither the vocabulary nor the
relationship questions contributed to the separation of the groups).

Only one statistically significant discriminant function, based upon the
level of information variables, was extracted from this analysis. This
function yielded a chi-square value of 6.39, with 2 degrees of freedom, sig-
nificant at $p < .05$. This function accounted for 20% of the variance in high
information questions and 29% of the variance in low level information ques-
tions. The function had a canonical correlation of .35 with group member-
ship.

A classification analysis based on pooled covariance matrix and individ-
ual discriminant scores was then used to test the effectiveness of the func-
tion obtained. The use of this discriminant function led to the correct
classification of 65% of the rehearsal subjects, or a modest 15% more than
could be classified correctly on the basis of chance.

Limitations

This study has several limitations. First, the amount of variance expla-
nation attributable to information types was attenuated by the low internal
consistencies of the various dependent measures. More reliable tests might
permit a more powerful discriminant function to be derived. A clearer pic-
ture of the specific outcomes of these study strategies might be obtained
through the use of longer tests or tests that evaluate the learning of fewer
information categories. This might permit sharper distinctions to be drawn
between the strategies compared.
Second, this study asked subjects to study only a short text (2000 words) drawn from a single content area. It is quite possible that the demands of different types of text or information would have resulted in different findings. Future research needs to examine these procedures and information classifications with a variety of materials, and under a variety of conditions.

Third, this study found differences in the two study procedures in specific types of learning. However, it should be noted that this study measured learning on the basis of a recognition (i.e., multiple-choice test) paradigm only. Very different results might have been obtained using a recall format. Future research needs to examine a variety of strategies and information types.

Conclusions

This study examined the possibility that two study strategies would lead to structurally different learning outcomes. Subjects in the two study groups were able to answer significantly more questions about cell theory than were their counterparts in the read-only or test-only groups. This indicates that the two study techniques led to higher performance than would be expected on the basis of either prior knowledge or a reading-only strategy. Also, the study group subjects performed better on the text information which they studied than on the information which they had only read. These findings indicate that the GPO and vocabulary rehearsal strategies examined here are effective study tools.
The two study groups scored equally well on the test of learning used in this study. This indicates that the amounts of learning derived from these techniques, given equal study times are substantially the same. However, the discriminant analysis indicated that types of information that contributed to these overall scores were not the same across groups. The vocabulary strategy led to superior performance on the low level or subordinate information questions, while the GPO strategy led to somewhat better performance on the high level or superordinate information questions. Although one group defined vocabulary and the other attempted to construct relationships between these same vocabulary terms, groups performed equally well on the vocabulary-relationship questions. (This might be due to the limitations imposed by low reliability tests or it could be due to the fact that these question types have no psychological reality.)

This study, together with previous research (Dee-Lucas & DiVesta, 1980), indicate that study strategy training should attempt to teach students how to use several rehearsal strategies. Such training should do more than make students proficient in the mechanics of these techniques, however. The training should also make students aware of the specific outcomes which might be derived from each strategy. With such strategy-outcome information as part of their metacognitive armamentarium, learners could then use the strategies in combination, or they could be selected from on the basis of desired outcomes.
References


Brady, P.J., & Rickards, J.P. How personal evaluation affects the underlining and recall of prose material. *Journal of Reading Behavior*, 1979, 11, 61-68.


Riley, J.D., & Dyer, J. The effects of notetaking while reading or listening. Reading World, 1979, 19, 51-56.


Acknowledgement

This research was supported by the Research Board, Graduate College, University of Illinois at Chicago Circle.
Table I. Mean proportions and standard deviations for test performance of the vocabulary rehearsal group (n=26), the graphic post organizer rehearsal group (n=25), the read-only group (n=21), and the test-only group (n=44).

<table>
<thead>
<tr>
<th>Dependent Measure</th>
<th>Groups I (Vocabulary) M/SD</th>
<th>Groups II (GPO) M/SD</th>
<th>Groups III (Read Only) M/SD</th>
<th>Groups IV (Test Only) M/SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studied Items</td>
<td>.71/.15</td>
<td>.69/.10</td>
<td>.61/.15</td>
<td>.50/.14</td>
</tr>
<tr>
<td>Non-Studied Items</td>
<td>.58/.21</td>
<td>.57/.20</td>
<td>.54/.16</td>
<td>.45/.16</td>
</tr>
<tr>
<td>High Level-Studied</td>
<td>.80/.14</td>
<td>.84/.09</td>
<td>.69/.18</td>
<td>.57/.20</td>
</tr>
<tr>
<td>Low Level-Studied</td>
<td>.63/.19</td>
<td>.54/.17</td>
<td>.53/.21</td>
<td>.43/.16</td>
</tr>
<tr>
<td>Vocabulary-Studied</td>
<td>.77/.13</td>
<td>.75/.12</td>
<td>.62/.18</td>
<td>.52/.17</td>
</tr>
</tbody>
</table>
Table II: Univariate F-Ratios derived from comparisons of vocabulary and graphic post organizer rehearsal groups on dependent measures. (df 1, 49).

<table>
<thead>
<tr>
<th>Measure</th>
<th>Univariate F-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Test Score</td>
<td>1.48</td>
</tr>
<tr>
<td>Non-Studied Items</td>
<td>0.43</td>
</tr>
<tr>
<td>High Level-Studied</td>
<td>1.40</td>
</tr>
<tr>
<td>Low Level-Studied</td>
<td>2.90</td>
</tr>
<tr>
<td>Vocabulary-Studied</td>
<td>0.55</td>
</tr>
<tr>
<td>Relationship-Studied</td>
<td>0.23</td>
</tr>
</tbody>
</table>