The purpose of this study was to determine the effects of note taking and review on retention of information presented by lecture. One hundred seventy-two undergraduates enrolled in an introductory psychology course served as subjects for the experiment. All subjects listened to a lecture while engaging in study strategies consisting of combinations of note taking and review. Performance was measured immediately and/or one week later by free recall and by verbatim and paraphrase completion tests. Results showed that taking and reviewing notes yeilded maximum retention, while listening only, without review, resulted in poorest performance. The benefit of note taking appears to be derived from having a subsequent opportunity to review notes, and not from the act of note taking itself. Encoding differences as a function of note taking appear to be minimal, while the external storage function is of primary importance. (Author/TS)
THE EFFECTS OF NOTE-TAKING AND REVIEW ON RETENTION OF INFORMATION PRESENTED BY LECTURE

Nicholas H. Van Matre and John F. Carter

Navy Personnel Research and Development Center
San Diego, California 92152

Washington, D.C., 1975
Abstract

Subjects listened to a lecture while engaging in study strategies consisting of combinations of note-taking and review. Performance was measured immediately and/or one week later by free recall, and verbatim and paraphrase completion tests. Taking and reviewing notes yielded maximum retention, while listening-only without review resulted in poorest performance. The benefit of note-taking appears to be derived from having an opportunity to subsequently review notes, and not from the act of note-taking itself. Encoding differences as a function of note-taking appear to be minimal, while the external storage function is of primary importance.
This study was conducted to investigate the effects of note-taking and review activities on the retention of information presented by lecture. While the effects of review have generally been found to be facilitative, the findings with respect to note-taking have been mixed. Eisner and Rhode (1959) and Berliner (1969) found no clearly beneficial effects due to note-taking, while Peters (1972) found that note-taking actually hindered performance. Other investigators have reported data supporting note-taking strategies as being effective for improving recall. (e.g., Crawford, 1925; DiVesta & Gray, 1972, DiVesta & Gray, 1973). Throughout these studies there has been an interest in discovering the functions served by note-taking. One hypothesis is that note-taking enhances encoding of the to-be-learned material into a subjectively more meaningful form. A second hypothesis is that notes serve only an external storage function, providing a means for the later retrieval of information for additional study. Fisher and Harris (1973) predicted greater importance for the encoding than for the external storage function, but found that the external storage function provided by notes was more beneficial to recall than the encoding function of the note-taking behavior itself.

The present investigation attempted to further clarify the role of note-taking and review in a paradigm designed to separate the encoding and external storage functions for immediate and long-term retention. The timing and opportunity for review were varied for note-taking and listen only groups. Both immediate and delayed tests were given with the opportunity for review always coming immediately before the test. This was done to conform to what we believe constitutes a strong test of the
external storage and encoding hypotheses and to more closely parallel the conditions which occur when notes are taken and reviewed in a naturalistic setting.

An additional variable involved the type of item appearing on a fill-in test given to the Ss. The encoding hypothesis assumes that Ss taking notes engage in more meaningful processing of the lecture material. If so, this ought to effect performance on a test requiring semantic processing of the material. A fill-in test where the stem is a paraphrase of the relevant portion of the passage fulfills this requirement (Anderson, 1972). Therefore, verbatim and paraphrase test items were administered in order to detect encoding differences between groups. It was predicted that where encoding was inferior, the group would perform worse on paraphrase items than verbatim items.

Method

Subjects

One hundred and seventy-two undergraduates enrolled in an Introductory Psychology course served as subjects in this experiment. Their participation partially satisfied a course requirement.

Design and Procedures

The Ss engaged in one of four study strategies while listening to a 17-minute recorded lecture. In one treatment subjects took notes during the lecture and then reviewed those notes for a five minute period following the lecture (Notes/Notes Review). In another condition subjects took notes, but engaged in a review without access to their notes (Notes/Mental Review).
In the third, subjects listened, but took no notes during the lecture, and then engaged in a mental review of what they had heard (Listen/Mental Review).

In the fourth condition, Ss listened to the lecture and then engaged in a letter cancellation task during the review period (Listen No Review). Orthogonal to these treatments was Review/Test Interval as a between-subject factor involving either immediate review and testing, or review and testing after a one week delay; and two levels of verbal ability. In addition, in order to compare delayed test performance for subjects who took the immediate test with that for subjects who did not, the subjects who took the immediate test were asked to return for the delayed test. These subjects, however, were not given a second review period.

Immediately following the review period, free recall, and verbatim and paraphrase completion tests were administered. The free recall test consisted of having subjects write everything they could remember about the lecture. This was scored in terms of the number of pre-defined information units recalled. The completion test consisted of 30 verbatim and 30 paraphrase fill-in items. A verbatim item was a verbatim reproduction of a sentence contained in the lecture with a substantive word or number deleted. Paraphrase equivalents were written for each of these to yield two test forms. Each item appeared on both forms, as a verbatim item on one form, and as a paraphrase item on the other form. The two forms were administered within each experimental group in a counterbalanced fashion.
Results and Discussion

Data from the free recall test were submitted to a 4 x 2 x 2 analysis of variance involving as factors: Study Strategy (Notes/Notes Review vs. Notes/Mental Review vs. Listen/Mental Review vs. Listen/No Review), Review/Test Interval (immediate vs. one-week delay), and Verbal Ability (high vs. low). Means for this analysis are shown in Table 1 collapsing across Verbal Ability which was of no interest since there was no interaction with any other variable. All main effects were significant, indicating superior recall of information units for the high ability subjects, and on the immediate test (p < .05). The Notes/Notes Review condition produced reliably better performance than the two mental review groups, which were in turn significantly better than the No-Review group (p < .01).

The differences among study strategies indicated that there was no beneficial effect to be derived from the note-taking act itself unless it was combined with a subsequent review of the notes. These findings clearly supported the external storage hypothesis for free recall performance.

An analysis of the ratio of words written to information units recalled was run. Results paralleled those for free recall indicating that one way in which note-taking with notes review influenced recall was to decrease the amount of irrelevant material output by Ss.

Completion test data were analyzed by a 4 x 2 x 2 x 2 analysis of variance employing the same factors as before, plus Item Type (verbatim vs. paraphrase) as a within-subject factor. Means for this analysis are
presented in Table 2. As in the former analyses, all main effects were significant \( p < .05 \) including the Item Type effect. However, in addition, the three way interaction between Study Strategy, Review/Test Interval and Item Type reached significance \( F = 3.62, \text{df} = 3/156, p < .025 \). This can best be described as follows. For the immediate test, performance was better on verbatim than paraphrase items for all study strategies except the Notes/Notes Review condition, where there was no difference. However, on the delayed test, performance was better on verbatim items only for the Notes/Notes Review condition. We think this finding is interpretable in the following way. On an immediate test verbatim performance ought to be better than paraphrase performance unless conditions are present which facilitate semantic encoding (Anderson, & Biddle, 1975). The encoding hypothesis would predict this for both note-taking conditions, while the external storage hypothesis would do so only for the Notes/Notes Review condition. Following a delay on the other hand, differences between verbatim and paraphrase performances tend to diminish, most likely as a result of forgetting of superficially processed information. Only if some intervening event re-exposed subjects to a verbatim representation of the communication would greater verbatim performance be predicted. Again, the Notes/Notes Review condition represented such a case since an examination of subjects' notes indicated that they consisted largely of verbatim excerpts of lecture material.

Additional weight is added to this interpretation when the delayed test performance of the subjects who took the immediate test is considered. These subjects were given a review period immediately before
taking the immediate test, and then returned one week later for the delayed test, but without the opportunity for further review. If our reasoning is correct, there should not be a significant difference between performance on verbatim and paraphrase items for any of these groups, since there was no intervening event which should have reminded subjects of the verbatim expressions in the lecture. As predicted, the analysis of these data showed no significant differences between verbatim and paraphrase performance, either as a main effect or in interaction with the Study Strategy factor ($p > 7.05$). Taken together these data give strong support to the external storage hypothesis, while failing to support the notion that the act of note-taking facilitates encoding.

A question tangential to the primary purpose of this study was whether taking an immediate facilitated free recall performance. Although the facilitating effects of taking an immediate test are well documented (e.g., Roderick & Anderson, 1968; Spitzer, 1939), the typical dependent measure in previous research has been some sort of cued performance. In this study the analysis indicated superior delayed-test performance for subjects taking the immediate free recall test ($\bar{X} = 53.7$) over subjects who took only the delayed test ($\bar{X} = 37.2$) ($F = 25.98$, $df = 1/143$, $p < .001$). Thus, the effect in this study occurred despite the fact that the group that took the immediate test had no opportunity for review before taking the delayed test as did the delayed-test group (although they had reviewed before taking the immediate test). The effect of an immediate test, then, evidently holds for free recall performance over a retention interval of one week.
In conclusion, the most important finding from this study is the strong support given to the external storage hypothesis, and the general lack of support for the encoding hypothesis. These data indicate that it is better not to take notes at all than to take them and not review them prior to a test. This conclusion holds whether the criterion performance involves cued or uncued testing. It appears from these data and those of Fisher and Harris (1973) that the encoding hypothesis can largely be discounted as an important consideration in instructional situations involving note-taking.
<table>
<thead>
<tr>
<th>Study Strategy</th>
<th>Immediate</th>
<th>Delayed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notes/Notes Review</td>
<td>88.8</td>
<td>53.6</td>
</tr>
<tr>
<td>Notes/Mental Review</td>
<td>61.3</td>
<td>31.8</td>
</tr>
<tr>
<td>Listen/Mental Review</td>
<td>63.6</td>
<td>33.0</td>
</tr>
<tr>
<td>Listen/No Review</td>
<td>56.5</td>
<td>21.6</td>
</tr>
</tbody>
</table>
Table 2

Mean Completion Test Performance

<table>
<thead>
<tr>
<th>Study Strategy</th>
<th>Item Type</th>
<th>Immediate</th>
<th>Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notes/Notes Review</td>
<td>Verbatim</td>
<td>17.7</td>
<td>13.4</td>
</tr>
<tr>
<td></td>
<td>Paraphrase</td>
<td>17.9</td>
<td>11.2</td>
</tr>
<tr>
<td>Notes/Mental Review</td>
<td>Verbatim</td>
<td>14.2</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td>Paraphrase</td>
<td>13.1</td>
<td>7.3</td>
</tr>
<tr>
<td>Listen/Mental Review</td>
<td>Verbatim</td>
<td>15.6</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td>Paraphrase</td>
<td>13.8</td>
<td>8.3</td>
</tr>
<tr>
<td>Listen/No Review</td>
<td>Verbatim</td>
<td>16.2</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>Paraphrase</td>
<td>14.4</td>
<td>7.6</td>
</tr>
</tbody>
</table>
References


Berliner, D. C. Effects of test-like events and note-taking on learning from lecture instruction. Proceedings, 77th Annual Convention, APA, 1969


Eisner, S. & Rohde, K. Note-taking during or after the lecture. Journal of Educational Psychology. 1959, 50, 301-304.


