Evidence has accumulated indicating that high arousal or activation facilitates long-term retention (LTR) but depresses short-term retention (STR) relative to low activation in list learning. The present study extended this research to the learning and retention of text by children. It specifically investigated the effects of questions inserted into prose material on STR and LTR. Two hundred and ninety third to sixth grade students read two passages. Subjects were randomly assigned to conditions and answered questions inserted in both, either, or neither passage. Grades five through six read more complex versions of the passages. Subjects completed a 25-item multiple-choice retention test immediately (STR) and one week later (LTR). An ANOVA and post-hoc pair-wise comparisons revealed the group answering questions in both passages scored significantly higher on both STR and LTR than controls, indicating a facilitative effect of adjunct questions. A text complexity and learner strata interaction was suggested. The complexity of such activation research with text was discussed in terms of the conditions of testing and the nature of questions. (Author)
Technical Report No. 201

CHILDREN'S PROCESSING OF PROSE: THE EFFECTS OF QUESTION AROUSAL, TEXT COMPLEXITY, AND LEARNER STRATA ON SHORT- AND LONG-TERM RETENTION

by

Frank H. Farley and Roger R. Eischens

Report from the Project on Motivation and Individual Differences in Learning and Retention

Frank H. Farley, Principal Investigator

Wisconsin Research and Development Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

December 1971
Statement of Focus

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Motivation and Individual Differences in Learning and Retention Project from Program 1. General objectives of the Program are to generate new knowledge about concept learning and cognitive skills, to synthesize existing knowledge, and to develop educational materials suggested by the prior activities. Contributing to these Program objectives, the Learning and Memory Project has the long-term goal of developing a theory of individual differences and motivation. The intermediate objective is to generate new knowledge of the learning and memory processes, particularly their developmental relationship to individual differences and to motivation.
Contents

List of Tables v
Abstract vii
Introduction and Review of the Literature 1
Method 3
Results 5
Discussion 9
References 11
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Distribution of Subjects by Grade</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Treatments According to the Number of Questions</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Mean Retention Scores for 3rd Graders by Groups for both the Short- and Long-term Tests</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>Summary of Scheffé's Post Hoc Pair-Wise Comparison of LTR-E (Grade 3)</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>Mean Retention Scores for Grades 4, 5, and 6 Collapsed</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>Mean Retention Scores by Groups for both Short- and Long-term Tests at Grades 4, 5, and 6</td>
<td>6</td>
</tr>
</tbody>
</table>
Abstract

Evidence has accumulated indicating that high arousal or activation facilitates long-term retention (LTR) but depresses short-term retention (STR) relative to low activation in list learning.

The present study extended this research to the learning and retention of text by children. It specifically investigated the effects of questions inserted into prose material on STR and LTR.

Two hundred and ninety grade 3-6 students read two passages. Subjects were randomly assigned to conditions and answered questions inserted in both, either, or neither passage. Grades 5-6 read more complex versions of the passages. Subjects completed a 25-item multiple-choice retention test immediately (STR) and one week later (LTR).

An ANOVA and post-hoc pair-wise comparisons revealed the group answering questions in both passages scored significantly higher on both STR and LTR than controls, indicating a facilitative effect of adjunct questions. A text complexity and learner strata interaction was suggested. The complexity of such activation research with text was discussed in terms of the conditions of testing and the nature of questions. Important issues requiring further examination were outlined.
Research on the relationship of arousal to memory consolidation has suggested that arousal may differentially affect short- and long-term retention. It is hypothesized that learning under high arousal produces a more actively consolidating trace that is relatively unavailable for short-term recall but is ultimately better consolidated for long-term recall than learning under low arousal (Farley, in press).

Kleinsmith and Kaplan (1963, 1964) have reported evidence supporting the differential effect of arousal on short- and long-term memory. In the first study (Kleinsmith & Kaplan, 1963) words and numbers were used as stimuli and responses, respectively, in paired-associates (PA) learning with arousal defined in terms of galvanic skin response (GSR) deflections to the words during learning. High-arousal words were then separated in the recall analysis from low-arousal words. Five recall intervals were used: immediate, 20 minutes, 45 minutes, 1 day, and 1 week. Responses learned under low arousal were recalled better at the immediate test, were undifferentiated at the 20-minute test from responses learned under high arousal, and showed "classical forgetting" at all the remaining test intervals, whereas on the latter tests the high-arousal material demonstrated reminiscence. Kleinsmith and Kaplan (1964) successfully replicated this finding using 0% association value nonsense syllables rather than words, and three retention intervals: immediate, 20 minutes, and 1 week.

Walker and Tarte (1963), using homogeneous and mixed lists of high and low arousal words, have obtained results comparable to those of the Kleinsmith and Kaplan studies. Farley (1969) has recently used the stimulus words of the Walker and Tarte (1963) study in a free learning experiment. He obtained results similar to those of Walker and Tarte (1963) and Kleinsmith and Kaplan (1963, 1964) with respect to the long-term recall measure, but he did not obtain the crossover effect between immediate and long-term recall. However, in other studies, Farley and Gilbert (1970), Lovejoy and Farley (1969), Manske and Farley (1971) and Osborne and Farley (1971) have obtained the expected crossover effect.

Other experiments have attempted to manipulate general arousal level rather than stimulus-specific arousal. Alper (1948) attempted to induce arousal by administering "ego-oriented instruction." She tested for recall immediately after learning and one day later. Ego-oriented Ss not only recalled significantly more new items on Day 2 than on Day 1 but also recalled on Day 2 significantly more of the same items they had recalled on Day 1 than did the "task-oriented" Ss.

King and his associates (Harper & King, 1967; King & Dodge, 1965; King & Walker, 1965) have used a method of delayed auditory feedback (DAF) to induce arousal. They found that immediate retention of prose material practiced under DAF of .2 to .8 seconds is significantly poorer than that obtained from appropriate controls. However, on a long-term retention test, material practiced under DAF yielded greater retention relative to the initial amount of material recalled in comparison to the control group.

Berlyne, Borsa, Craw, Gelman, and Mandell (1965) and Berlyne, Borsa, Hamacher, and Koenig (1966) have induced arousal by using white noise. They found that white noise during presentation of stimulus and response terms in training trials significantly increased recall in a test trial given 24 hours later. They also found that during training on Day 1, white noise under all presentation conditions had no detrimental effect on immediate recall.

Thus, the bulk of the foregoing studies...
employing arousal-producing stimulus terms, frustrating tasks, and white noise suggest that arousal facilitates long-term recall, while the results concerning the effect of arousal upon immediate recall have been less certain.

The present study attempts to investigate the role of activation in the retention of prose material. The specific question to be investigated is whether a general rise in arousal level induced by questions inserted in text will differentially affect short- and long-term retention. In an earlier study Natkin and Stahler (1969) reported that questions inserted in text can produce the crossover effect predicted by the arousal hypothesis of Walker and Tartre (1963). Natkin and Stahler (1969) posited that asking the S a question produces an increase in arousal and that this effect can be controlled by use of a second passage read prior to the learning material in which the Ss are exposed to questions. It was assumed that any arousal effects produced by a stimulus would habituate with frequent stimulus exposure. Natkin and Stahler predicted that Ss asked to answer questions inserted into learning material, who were not exposed to questions in a passage read before the learning material, would recall less on a short-term retention test than Ss who either had questions in the pre-exposure material plus learning material or who did not have questions in the learning material. They also predicted that the Ss in the high-arousal conditions (questions in learning material, none in pre-exposure material) would do significantly better on a long-term retention test. Natkin and Stahler's (1969) results supported the predictions concerning the effects of questions and the authors attributed this to arousal-inducing effects of the questions.

Earlier evidence that inserting questions into text will facilitate retention of material that is not question specific has been reported by Hershberger and Terry (1965), Rothkopf (1966), Rothkopf and Bisbicos (1967), and Frase (1967, 1968). This research has revealed two important results: (a) questions facilitate learning even without knowledge of results, and (b) questions have a general facilitative effect as well as a specific facilitation upon question-specific information. These effects have been explained by Rothkopf (1965) and Frase (1968) in terms of the "mathemagenic" hypothesis. This analysis stresses that the acquisition and retention of information from printed material can be related to a variety of ongoing responses, summarized by the term "mathemagenic"—they give rise to learning. These responses can be brought under control of test-like events (such as questions) which occur in conjunction with the reading materials. In the above series of studies the groups who have received questions in the material have in almost all cases performed better than the controls on short-term retention tests, contrary to the Natkin and Stahler (1969) results. The following study was initiated because of the inconsistency of the research reported concerning questions, as well as the desire to extend the present first author's investigations of the effects of activation in list learning to learning from text, and to investigate the prose learning of children. The experiment was designed to study the effects of putative increases in arousal induced by adjunct questions inserted into text upon short- and long-term retention of non-question-specific information within prose material. An experimental procedure similar to that of Natkin and Stahler (1969) was employed with necessary adaptations due to the change in the age level of the sample. The specific hypothesis to be tested was that increased arousal due to adjunct questions inhibits short-term retention but facilitates long-term retention.
Sample

The Ss were 342 elementary school children from Sussex, Wisconsin. Table 1 gives the distribution of the Ss by grade.

Table 1. Distribution of Subjects by Grade

<table>
<thead>
<tr>
<th>Grade</th>
<th>Number of Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>101</td>
</tr>
<tr>
<td>4</td>
<td>80</td>
</tr>
<tr>
<td>5</td>
<td>86</td>
</tr>
<tr>
<td>6</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong> 342</td>
</tr>
</tbody>
</table>

An attempt was made to test all the Ss in these grades in one school.

Materials

The materials for the study were two passages of approximately 500 words each. The first passage was about the plants and animal life in the swamps of the Everglades. The second one contained information about life styles of two forest animals, the sloth and the wild pig. The first passage served to control the habituation reaction to questions and the second one was used as the learning material. Two versions of each passage were prepared. Version A had short, simple sentences and was administered to the 3rd and 4th grades. Version B, administered to the 5th and 6th grades, had the same content but with more complex sentence structure. The materials were prepared with the help of reading experts and teachers to insure a grade-related level of complexity of content and vocabulary. For the 4th, 5th and 6th grades two forms of each version were stapled separately into two booklets. The first form had experimental questions and the second one none. The swamp material (pre-exposure material) contained nine adjunct questions inserted approximately every 55 words. The learning material had six adjunct questions, i.e., one question after every 85 words. The experimental questions were short-answer type covering one detail of the information presented on the previous page. Each question to be inserted was randomly selected out of four questions originally prepared. The 16 questions not selected as experimental (adjunct) questions for the learning passage were combined with seven additional questions to form a 25-item criterion or outcome (retention) measure. Each question on the criterion test was presented as a four-distractor multiple choice item. None of the items on the post-test overlapped with those used in the text, and the items covered different details than the experimental (adjunct) questions.

The procedure was slightly different for the 3rd grade in that there were four forms of each version of the learning material. The basic booklet consisted of six pages of approximately 85 words each. The other three booklets had questions inserted after every 85 words, 170 words, and 255 words. The adjunct questions were always on a separate page inserted as another page in the text. The pre-exposure material was the same as for the 4th, 5th, and 6th grades (i.e., containing either zero or nine adjunct questions).

Design

The 3rd grade had eight groups of Ss differing in the number of questions they received. The questions were the results of various com-
Table 2. Treatments According to the Number of Questions

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-exposure Material</th>
<th>Learning Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>8 (Control Group)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Combinations of the versions of both the passages. The treatments according to number of questions in each booklet are given in Table 2.

All the Ss received an immediate recall test and a long-term recall test (one week), both of which were administered at the same time of day and by the same E.

The 4th, 5th, and 6th grades were randomly divided into two groups each, one receiving the questions and the other not for each passage (i.e., the pre-exposure and the learning passages). Similar to above, each S received two tests of retention.

Procedure

Testing for the 3rd grade required randomly dividing the class into eight groups. The materials were first randomly distributed and then E read a standardized set of instructions, after which the Ss were asked to begin reading. The Ss were instructed to continue reading on to the second booklet after reading the pre-exposure material. A short-term retention test followed the learning material; a long-term retention test was administered one week later. In addition, a retention test covering the pre-exposure material was given after the long-term retention test of the learning material.

Testing for the remaining three grades was similar to the above except for differences in the number of groups in line with the research design outlined earlier.
The 3rd grade data were analyzed separately from the data for grades 4, 5, and 6. Table 3 provides the mean retention scores for the 3rd graders in the various groups.

The data from the three retention measures were analyzed employing a univariate-multivariate analysis of variance technique. The multivariate F-ratio for STR-L, LTR-L was not significant ($F = .8253$, $df = 14/184$, $p < .6410$). The univariate F-ratios for STR-L and LTR-L were STR-L: $F = .7689$, $df = 14/184$, $p < .6410$; LTR-L: $F = 1.0381$, $df = 14/184$, $p < .4102$.

Further analysis indicated that although the univariate test for differences between STR-L and LTR-L was significant ($F = 22.8793$, $df = 1/93$, $p < .0001$), the univariate test for treatment effects on the difference scores was not significant ($F = .7393$, $df = 7/93$, $p < .6392$).

The analysis of the retention scores of the pre-exposure material indicated a significant main effect for the experimental treatment ($F = 3.8945$, $df = 7/93$, $p < .001$). Scheffé's test for post hoc pair-wise comparisons was utilized to find which experimental groups were significantly different from each other on the long-term retention measure. Table 4 presents the results of this analysis. Scheffé's test indicated that Group 2, the group which answered nine questions in the first booklet and three questions in the second booklet, scored significantly better than both Groups 6 and 7 on the retention test. Group 6 answered zero questions in the first booklet and two questions in the second booklet, while Group 7 answered nine questions in the first booklet and zero questions in the second booklet. None of the remaining mean differences were significant at the .05 level or better.

Where grades 4, 5, and 6 are concerned, Table 5 presents the mean number of correct responses on the two retention measures (immediate and one week) for the experimental conditions independent of grades.

A univariate and multivariate analysis of variance revealed significant effects due to experimental treatments. The multivariate test yielded $F = 1.811$, $df = 6/456$, $p < .09$.

### Table 3. Mean Retention Scores for 3rd Graders by Groups for both the Short- and Long-term Tests

<table>
<thead>
<tr>
<th>Group</th>
<th>Learning Material</th>
<th>Pre-exposure Material</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STR-L</td>
<td>LTR-L</td>
</tr>
<tr>
<td>1</td>
<td>14.25</td>
<td>12.17</td>
</tr>
<tr>
<td>2</td>
<td>14.75</td>
<td>13.00</td>
</tr>
<tr>
<td>3</td>
<td>14.31</td>
<td>13.62</td>
</tr>
<tr>
<td>4</td>
<td>12.55</td>
<td>10.36</td>
</tr>
<tr>
<td>5</td>
<td>15.00</td>
<td>13.64</td>
</tr>
<tr>
<td>6</td>
<td>13.46</td>
<td>13.23</td>
</tr>
<tr>
<td>7</td>
<td>13.31</td>
<td>11.62</td>
</tr>
<tr>
<td>8</td>
<td>11.84</td>
<td>10.46</td>
</tr>
</tbody>
</table>
Table 4. Summary of Scheffé's Post Hoc Pair-Wise Comparison of LTR-L (Grade 3)

<table>
<thead>
<tr>
<th>Group (Rank Order)</th>
<th>2</th>
<th>5</th>
<th>8</th>
<th>3</th>
<th>1</th>
<th>4</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>2(9.17)</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>p &lt; .05</td>
<td>p &lt; .025</td>
<td></td>
</tr>
<tr>
<td>5(8.14)</td>
<td>-</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>8(7.92)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>3(7.08)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>1(6.25)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>4(6.18)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>6(5.92)</td>
<td>p &lt; .05</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>7(5.69)</td>
<td>p &lt; .025</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5. Mean Retention Scores for Grades 4, 5, and 6 Collapsed

<table>
<thead>
<tr>
<th>Groups</th>
<th>Treatment</th>
<th>STR-L</th>
<th>LTR-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>QQ</td>
<td>16.30</td>
<td>14.67</td>
</tr>
<tr>
<td>2</td>
<td>NQ</td>
<td>15.05</td>
<td>12.97</td>
</tr>
<tr>
<td>3</td>
<td>QN</td>
<td>14.98</td>
<td>13.58</td>
</tr>
<tr>
<td>4</td>
<td>NN</td>
<td>13.71</td>
<td>12.28</td>
</tr>
</tbody>
</table>

Table 6. Mean Retention Scores by Groups for Both Short- and Long-term Tests at Grades 4, 5, and 6

<table>
<thead>
<tr>
<th>Grades</th>
<th>Group</th>
<th>Treatment</th>
<th>STR-L</th>
<th>LTR-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>QQ</td>
<td>16.26</td>
<td>14.47</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>NQ</td>
<td>16.18</td>
<td>12.81</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>QN</td>
<td>15.63</td>
<td>14.05</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>NN</td>
<td>13.55</td>
<td>12.45</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>QQ</td>
<td>15.71</td>
<td>14.38</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>NQ</td>
<td>13.81</td>
<td>12.29</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>QN</td>
<td>15.23</td>
<td>14.09</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>NN</td>
<td>11.89</td>
<td>10.63</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>QQ</td>
<td>17.00</td>
<td>15.19</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>NQ</td>
<td>15.11</td>
<td>13.94</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>QN</td>
<td>13.88</td>
<td>12.31</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>NN</td>
<td>15.60</td>
<td>13.70</td>
</tr>
</tbody>
</table>

The univariate Fs for STR-L, LTR-L, and STR-L + LTR-L were significant (STR-L: F = 2.852, df = 3/229, p < .038; LTR-L: F = 2.641, df = 3/229, p < .050; STR-L + LTR-L: F = 2.915, df = 3/229, p < .035). This indicates that there is a significant main effect for treatments. A post hoc analysis was undertaken employing Scheffé's test for pair-
wise comparisons. The results of this analysis indicated that Group 1 had performed significantly better on both the short- and long-term measures than the control Group 4 (STR-L: $F = 10.49, df = 3/229, p < .02$; LTR-L: $F = 8.83, df = 3/229, p < .05$), while all other contrasts were not significant.

The ANOVA indicated no significant differences between grades and also no significant grade x treatment interaction.

A univariate analysis of the difference between STR-L and LTR-L indicated that a significant amount of forgetting had occurred over one week ($F = 84.31, df = 1/229, p < .0001$), but that no differences were present due either to grade or experimental treatment. Table 6 presents the retention scores by grade and group.
The present results did not verify the prediction that questions inserted into the second of two prose passages would induce a high level of arousal and cause inhibition of short-term retention test performance and facilitation of long-term performance when compared with a group having no questions and a group having questions in both passages.

The results did indicate that there existed a significant main effect for the experimental treatment and that the difference was the result of Group 1 (questions in both passages) scoring higher than the control group on both short- and long-term retention tests. In light of previous research by Rothkopf (1965) and Frase (1967, 1968) this is not a completely unexpected result, as will become apparent later.

A point that needs to be considered is that although the analysis indicated Group 1 to score significantly better than the control group (Group 4) when the results are looked at across grades, a close analysis of the means by grades (Table 6) reveals that most of these differences occur in 5th grade. Although the mean scores in grade 4 tend to appear in agreement with those of grade 5, in grade 6 the only agreement with the other grades is the position of Group 1 relative to the remaining groups.

At this juncture it should be noted that much of the ensuing discussion is not based on statistically significant findings but upon data which only approach traditional levels of significance.

One possibly important feature of the present data seems to be a trend indicated by the performances of similar groups across grade levels: grade 4 scores higher than grade 3, grade 6 groups score higher than similar groups in grade 5, with the partition at Group 3 on both retention measures. A seemingly important consideration is that groups in 4th grade score higher than the 5th grade groups on STR-L and equal or better on LTR-L. It might also be noted that the rank order of groups is the same for grades 3 and 5, and that grade 4 is more similar to grade 6 than grades 3 or 5 when one considers the relative position of groups. The appropriate question seems to be whether this pattern suggests some form of interaction between characteristics of the learner (grade) and the complexity level of the material which has an effect on retention. Rather than speculate what might cause this pattern to exist, it will only be suggested that further research might investigate this interaction between sentence complexity, experimental manipulation, and the information processing capacity of the learner.

Rothkopf (1965) and Frase (1969) have reported that an important factor to consider concerning the effects of adjunct questions upon nonadjunct question-specific retention is the amount of material between adjunct questions. If questions are inserted too frequently their generally facilitative effect disappears and they may even have an inhibiting influence. There is a strong possibility that this situation was present in this study, which would account for the nonsignificant results when other groups were compared with the control group. It would appear logical to assume that as the reading skill of the S increased the optimal amount of material between questions would increase, therefore causing these Ss to be inhibited by a question every 25 words, while for a S at a less efficient level of information processing, a question every 85 words might serve to facilitate incidental retention. This reasoning might possibly explain the performance of Group 4 in 6th grade. The control Group 4 scored from 3.0 to 3.5 points below Group 1 in the other three grades; however, in 6th grade the difference is only slightly over 1.0. In addition, Group 4 scored lower than
Groups 2 and 3 in grades 3, 4, and 5 but above these groups in grade 6.

Since the results tend to agree with those of Rothkopf (1965) and Frase (1967) rather than supporting an arousal x retention interval interaction hypothesis, the question as to why adjunct questions produce a general facilitative effect on retention remains unclear. An arousal hypothesis of question effects is not entirely disconfirmed, however, as there is reason to believe that in prose learning contexts arousal will have facilitative effects on both short- and long-term retention tests (Farley, 1971). Also, Farley (1969) has reported that when high-arousal word lists are compared to low-arousal word lists, performance is undifferentiated on an immediate test but the high-arousal words are better recalled on a long-term test, whereas when a mixed list (high- and low-arousal words) is used, high arousal words are better recalled at both retention intervals. This finding suggests that putative "arousal events" in prose (e.g., adjunct questions) within the "low-arousal context" of the prose passage will facilitate retention test performance at both short- and long-term intervals. Such an interpretation suggests parallels between the role of adjunct questions and the Von Restorff effect that might reward closer examination. It is also possible that the temporal parameters of short-term retention in the prose learning study are so vastly different from the one-trial list learning research that no short-term inhibitory or interference effects of arousal can be detected, because in the prose research both short- and long-term retention tests are tapping long-term memory. Thus, facilitative effects of arousal would be expected on both tests. Clearly the problem of studying short-term memory in connected discourse is complex and will require considerable revision and/or refinement of present approaches, with one promising possibility being the use of probes.

Some further basic questions need to be investigated: What types of cognitive processes might a question stimulate? What is the important function of the behavior following exposure to a question which can either facilitate or inhibit retention of non-question-specific information? How do the arousal, interest, topicality, and imagery-abstractness characteristics of the information interact with the type and frequency of questions to affect retention? Is there some interaction between reading efficiency, memory capacity, content characteristics, and spacing of questions which can account for the vast individual differences that exist in the data? May some of these individual differences be intrinsic (Jensen, 1967) to learning and memory processes or might they be extrinsic and identifiable in terms of intelligence, personality and so on? Would it be valuable to investigate the units of information in prose utilizing the same techniques as prior list learning studies (Maltzman, Kantor, & Langdon, 1966; Farley, 1969) concerned with the arousal-inducing power of individual units?
References

Alper, T. G. Task-orientation and ego-orientation as factors in reminiscence. Journal of Experimental Psychology, 1948, 38, 224-238.


Farley, F. H. Arousal, consolidation, and memory. Wisconsin Research and Development Center for Cognitive Learning, Theoretical Paper, in press.


Frase, L. T. Effect of question location, pacing and mode upon retention of prose material. Journal of Educational Psychology, 1968, 59, 297-301.


Harper, R. B., & King, D. J. Retention of material practiced under delayed auditory feedback: Preliminary observations on two additional variables. Psychological Reports, 1957, 21, 431-432.


Kleinsmith, L. J., & Kaplan, S. Paired-associate learning as a function of arousal


Maltzman, I., Kantor, W., & Langdon, B. Immediate and delayed retention, arousal, and the orienting and defense reflexes. Psychonomic Science, 1966, 6, 445-446.


National Evaluation Committee

Helen Bain
Immediate Past President
National Education Association

Lyle E. Bourne, Jr.
Institute for the Study of Intellectual Behavior
University of Colorado

Jeanne S. Chall
Graduate School of Education
Harvard University

Francis S. Chase
Department of Education
University of Chicago

George E. Dickson
College of Education
University of Toledo

Hugh J. Scott
Superintendent of Public Schools
District of Columbia

H. Craig Sipe
Department of Instruction
State University of New York

G. Wesley Sowards
Dean of Education
Florida International University

Benton J. Underwood
Department of Psychology
Northwestern University

Robert J. Wisner
Mathematics Department
New Mexico State University

Executive Committee

William R. Bush
Director of Program Planning and Management
and Deputy Director, R & D Center

Herbert J. Klausmeier, Committee Chairman
Director, R & D Center

Wayne Otto
Principal Investigator
R & D Center

Robert G. Petzold
Professor of Music
University of Wisconsin

Richard A. Rossmiller
Professor of Educational Administration
University of Wisconsin

James E. Walter
Coordinator of Program Planning
R & D Center

Russell S. Way, ex officio
Program Administrator, Title III ESEA
Wisconsin Department of Public Instruction

Faculty of Principal Investigators

Vernon L. Allen
Professor of Psychology

Frank H. Farley
Associate Professor
Educational Psychology

Marvin J. Fruth
Associate Professor
Educational Administration

John G. Harvey
Associate Professor
Mathematics

Frank H. Hooper
Associate Professor
Child Development

Herbert J. Klausmeier
Center Director
V. A. C. Henmon Professor
Educational Psychology

Stephen J. Knezevich
Professor
Educational Administration

Joel R. Levin
Associate Professor
Educational Psychology

L. Joseph Lins
Professor
Institutional Studies

Wayne Otto
Professor
Curriculum and Instruction

Thomas A. Romberg
Associate Professor
Curriculum and Instruction

Peter A. Schreiber
Assistant Professor
English

Richard L. Venezky
Associate Professor
Computer Science

Alan M. Voelker
Assistant Professor
Curriculum and Instruction

Larry M. Wilder
Assistant Professor
Communication Arts