A study examined the effects on comprehension of interspersing questions in text. The subjects, 103 seventh grade students in an urban school, were identified by scores on the California Achievement Test as either good or poor readers. They were then divided into four groups receiving the following treatments: passages without experimental questions, passages with questions but with instructions not to review the text, passages with questions and instructions to review, and passages with questions, instructions to review, and instructions to write out the responses. Following the reading of four content area passages, subjects took a multiple-choice comprehension test, which they retook two weeks later to determine the effects of the treatments on long term retention. The positive effects were noted for subjects in the question-review-write (QRW) treatment. The results indicated that interspersing question in text in the manner of QRW is effective for good readers on short term retention and short term and long term retention considered together, but not for long term retention. No significant differences appeared for poor readers in any of the treatment groups. The poor readers were reading materials at their frustration level, so even with the insertion of questions in text to reduce the amount of print encountered at one time, the readability level proved to be too difficult. (HTH)
The Effect of Interspersing Questions in Text: Evidence for "Slicing the Task"

Karen D. Wood
University of North Carolina at Charlotte

Running Head: Interspersed Questions

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Karen D. Wood

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."
The controversy over the effect of questions on learning from text, while temporarily put to rest, remains a dormant and still unsettled issue in the professional literature. Although research on the effects of interspersing questions in text is abundant, few, if any, classroom generalizations can be made from this area of experimentation. One reason for this continuing uncertainty can be attributed to the methods and procedures used to collect the data. In most instances, the studies were conducted under conditions which were not ecologically representative of the classroom environment. Another reason concerns the researchers' purposes for conducting the studies. Most were designed to test theories of learning and comprehending which were prevalent at the time. For example, during the behavioral era, the emphasis was on using questions to program and control student learning. These studies generally focused on manipulating certain variables such as question mode, question position and frequency of placement on the retention of intentional and incidental materials. Thus, this experimental approach has been termed a "variables orientation" to interspersed question research (Rickards and Denner, 1978, p. 313). In contrast, during the cognitive era, the focus was on varying "the depth of processing" (Craik and Lockhart, 1972, p. 675) required to answer the question. Through
this manipulation of question level, researchers have been able to make inferences regarding the processing characteristics effected by interspersing questions in text. As such, this approach has been termed a "process orientation" (Rickards and Denner, 1978, p. 313). Examining this issue from another perspective, that of the classroom practitioner, little is known about the effect questions may have on "slicing" or reducing the amount of print students encounter while reading their textbooks.

The concept of "slicing" was first described by Pearson and Johnson (1978) in conjunction with question-asking and later expanded by Readence and Moore (1980) as a means of re-examining the tasks required in reading text assignments and recasting them to match the abilities of the readers. At the sentence level, slicing can be accomplished by chunking the material into thought units to facilitate understanding (Carver, 1970). At the paragraph level, slicing is accomplished by interspersing questions in text to reduce the amount of print students must deal with at a given time.

To draw any conclusions from existing research about the benefits of using interspersed questions to "slice the task" would be difficult since, despite the recent influence of cognitive psychology, the interspersed question paradigm itself has remained much as it was in the sixties. Typically, one or two questions are inserted in a passage specially constructed to control for segment length and topical organization. The instructions stipulate that readers are not to turn back to a
Interspersed Questions - 3

page once it has been turned, nor to take any notes while reading (e.g., Hudgins, et al., 1979; Reynolds, Standiford, & Anderson, 1979; Rickards & Hatcher, 1978). Most frequently, college students have been employed as subjects in these studies (e.g., Andre et al., 1980; Boyd, 1974; Hiller, 1974; Koran & Koran, 1975; Rickards, 1976; Shavelson, Berliner, Ravitch & Loeding, 1974).

An obvious problem exists in the present studies in that they still possess the rudiments of programmed instruction. Alternating a page of text with a page containing a question is not a common classroom practice. In fact, it is difficult to determine if subjects actually attend to questions interspersed throughout the text, or, since often no overt response is required, if they instead tend to ignore the question entirely. Further, since no review of the text is permitted, the results of these studies are hardly generalizable to classroom situations which ordinarily permit review.

This artificiality of experimentation has been highly criticized in many reviews of the research on interspersed questions (Durkin, 1981; MacDonald-Ross, 1978; Rickards, 1979; Rickards & Denner, 1978). For example, Durkin's (1981) analysis of the literature identified several limitations and flaws which hinder the classroom applicability of the results of this research. First, she questioned the generalizability of the results to all students in that the subjects used were predominantly adults, most often easily accessible college students. Second, she doubted the general applicability of the
research to other types of school material since most typically
the material used was excerpted from psychology textbooks.
Third, she denounced the researchers for failing to describe
adequately their methodology.

Durkin recommended that the researchers broaden their scope
by investigating the effects of questions on students of all
ages, abilities, intelligence and socioeconomic backgrounds.
She emphasized the need to conduct the experiments in classroom
settings using actual school material and nonliteral level
questions. And, in an effort to add a new perspective to the
current research, she suggested that researchers in other
disciplines, such as reading education, should become involved
in the experimentation in cognitive psychology. It is to these
needs that the present study was directed.

Method

Subjects
The subjects for this study consisted of 103 seventh grade
students attending an urban school in the Southeast. The
students were identified as either good or poor readers using
scores on the California Achievement Test (1977). Poor readers
were defined as those students whose total reading grade equiva-
ient scores fell between 4.0 and 6.4 grade levels. Good readers
were defined as those students whose total reading grade equiva-
ient scores fell between 7.6 and 10.0 grade levels. The sub-
jects participating in the study consisted of 18 males and 24
females who met the definition of a good reader and 30 males
and 31 females who met the definition of a poor reader.
Experimental Materials

Passages. All subjects were required to read four content area passages. The passages, from a population of passages originally used in a study by Memory (1979), were randomly selected from six basal readers, six science texts, and six social studies texts on the Georgia State-Approved Textbook List. They were chosen because none exceeded the seventh grade level as determined by the Dale-Chall Readability Formula (1948).

Interspersed Questions. In this study, questions were placed after each text segment deemed conceptually significant and of appropriate length. This was done to approximate the manner in which a classroom teacher would determine question placement. The questions interspersed throughout the text passages were "meaningful learning postquestions," that is, "questions which require readers to organize paragraph details under the major concepts in the text segments" (Rickards and Hatcher, 1979, p. 543).

Criterion Tests. Following the reading of each passage, subjects took a multiple-choice test to assess their comprehension. Two weeks later, subjects retook each multiple-choice test to determine the effects of the treatments on long term retention. In order to include a broader sampling of comprehension tasks, the tests consisted of both textually explicit items which elicit information explicitly stated in the text and textually implicit items which elicit information not directly stated or implied in the text (Pearson and Johnson, 1978).
Judgments of Validity

A panel of judges was enlisted to determine the appropriateness of the following: 1) the content, placement and clarity of the questions interspersed in the text, and 2) the content and clarity of the multiple-choice items. The panel was composed of three doctoral students in reading education, all of whom were former seventh grade teachers. Items were accepted when two of the three judges reached agreement on their appropriateness.

Experimental Treatments

The experimental treatment conditions utilized in this study were as follows:

Treatment One (T₁) - No Questions (NQ). Subjects assigned to this group read text which contained no experimental questions.

Treatment Two (T₂) - Questions - No Review (QNR). Subjects assigned to this group read text containing the interspersed postquestions and instructions not to look back in the passage segment. There was one segment of text per page with the subsequent page containing only the postquestion. This treatment involved the experimental conditions most common to the interspersed question research to date.

Treatment Three (T₃) - Question - Review (QR). Subjects assigned to this group received the interspersed postquestions and instructions to look back in the passage segment for their response. Each page contained one segment of text and one postquestion. The page format described was intended to facilitate the review process.
Interspersed Questions - 7

Treatment Four (T₄) - Questions - Review - Write (QRW).

Subjects assigned to this group read text containing the interspersed postquestions, instructions to look back in the passage segment and instructions to write their responses on paper. Having the subjects write their responses was intended to force their attention to the questions in the text. One segment of text and one postquestion appeared on each page in addition to three lines on which to write the responses. Specifying the number of lines was designed to cue the reader to the length of response.

For each passage, four sets of booklets were constructed in the manner of the four treatment conditions. Specific instructions regarding the treatment were given on the front cover of the booklet. With the exception of Treatment Four (QRW) which required a written response, students were instructed not to write in the booklet.

Design

A 2 x 4 x 2 ANOVA with a repeated measure on the third factor was utilized in this study. The factor levels included achievement level (good and poor readers), treatment (the four conditions) and retention (short term and long term).

Using a Latin-square design, subjects were randomly assigned to one of four treatment arrangements (See Figure 1). This was done for three reasons: 1) to block any effects of treatment sequence; 2) to ensure that each subject encountered all four of the treatment conditions one time and 3) to block the effects of passage condition.
Figure 1

Description of the Four Treatment Arrangements to Which Subjects Were Randomly Assigned

<table>
<thead>
<tr>
<th>Subjects</th>
<th>NQ</th>
<th>QNR</th>
<th>QR</th>
<th>QRW</th>
</tr>
</thead>
<tbody>
<tr>
<td>NQ</td>
<td>QNR</td>
<td>QR</td>
<td>QRW</td>
<td>NQ</td>
</tr>
<tr>
<td>QNR</td>
<td>QR</td>
<td>QRW</td>
<td>NQ</td>
<td>QNR</td>
</tr>
<tr>
<td>QR</td>
<td>QRW</td>
<td>NQ</td>
<td>QNR</td>
<td>QR</td>
</tr>
<tr>
<td>QRW</td>
<td>NQ</td>
<td>QNR</td>
<td>QR</td>
<td>NQ</td>
</tr>
</tbody>
</table>
Procedures

The data for this study were collected in intact classrooms of 27 to 31 subjects. The data collection period for short term retention lasted six school days and the data collection period for long term retention lasted two school days. The researcher and two aides took over the classes for the duration of the study. Students were told that they were the participants in an attempt to find out the best way for them to gain information from their textbooks.

After introducing the purpose of the study, all subjects participated in a modeling activity designed to thoroughly acquaint them with the reading tasks and provide a natural transition during the actual experiment. Two weeks after the initial data collection period, the researcher returned to determine the effects of the treatments on long term retention.

Results

Table 1 presents the descriptive data for the investigation. Specifically, it shows the means and standard deviations for the two groups (good and poor readers) and for the four treatment conditions (NQ, QNR, QR, QRW) on both short term and long term retention.
Table 1
Summary of Cell Means and Standard Deviations for Good and Poor Readers on Short Term and Long Term Retention

<table>
<thead>
<tr>
<th></th>
<th>Good Readers</th>
<th>Poor Readers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>S.D.</td>
</tr>
<tr>
<td><strong>Short Term (ST)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NQ</td>
<td>5.88</td>
<td>2.13</td>
</tr>
<tr>
<td>QNR</td>
<td>6.75</td>
<td>1.89</td>
</tr>
<tr>
<td>QR</td>
<td>7.07</td>
<td>2.08</td>
</tr>
<tr>
<td>QRW</td>
<td>7.79</td>
<td>2.09</td>
</tr>
<tr>
<td><strong>Total (ST)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>M = 5.73</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Long Term (LT)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NQ</td>
<td>5.38</td>
<td>2.17</td>
</tr>
<tr>
<td>QNR</td>
<td>6.00</td>
<td>2.11</td>
</tr>
<tr>
<td>QR</td>
<td>6.26</td>
<td>2.37</td>
</tr>
<tr>
<td>QRW</td>
<td>6.50</td>
<td>2.29</td>
</tr>
<tr>
<td><strong>Total (LT)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>M = 4.98</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Total M</strong></td>
<td>6.45</td>
<td></td>
</tr>
</tbody>
</table>
The results of the 2 x 4 x 2 ANOVA displayed in Table 2 reveal statistically significant main effects for the three factors (achievement level, treatment and retention) manipulated in the experiment. The data revealed no significant interactions among these factors. With respect to achievement level, there was a statistically significant main effect, $F(1,101) = 94.34, p < .05$, with the retention scores of the good readers ($M = 6.45$) exceeding the scores of the poor readers ($M = 4.26$). With respect to retention, there was a statistically significant main effect, $F(1,101) = 62.48, p < .05$, with better scores on short term retention ($M = 5.73$) than on long term retention ($M = 4.98$).

Table 2 also shows a significant main effect for the treatment conditions across the two achievement levels $F(3,303) = 7.04, p < .05$. Scheffé's multiple comparison test was used to determine the means between which significant differences existed. The results indicated that the mean for the QRW treatment ($M = 11.11$) was significantly higher than the mean for the NQ treatment ($M = 9.28$) with no other significant differences between the treatment conditions.

<table>
<thead>
<tr>
<th>Achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td>There was a significant difference, $F(3,39) = 7.88, p &lt; .05$, between the retention scores of the good readers experiencing the various treatment conditions indicating that the treatments</td>
</tr>
<tr>
<td>Source of Variance</td>
</tr>
<tr>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Achievement</td>
</tr>
<tr>
<td>Error₁</td>
</tr>
<tr>
<td>Treatment</td>
</tr>
<tr>
<td>Treatment x Achievement</td>
</tr>
<tr>
<td>Error₂</td>
</tr>
<tr>
<td>Retention</td>
</tr>
<tr>
<td>Retention x Achievement</td>
</tr>
<tr>
<td>Error₃</td>
</tr>
<tr>
<td>Treatment x Retention</td>
</tr>
<tr>
<td>Treatment x Retention x Achievement</td>
</tr>
<tr>
<td>Error₄</td>
</tr>
</tbody>
</table>
Interspersed Questions - 10

had a significant effect on good readers' short term retention. The results of the Scheffe test revealed that the mean (M = 7.79) for treatment group QRW was significantly higher (p < .05) than the mean (M = 5.88) for treatment group NQ with no other differences reaching significance. However, with respect to the poor readers under study the treatments had no significant effect.

Treatment

In all instances where significant differences were obtained among the treatment conditions, the Scheffe multiple comparison test indicated that the QRW (questions - review - write) treatment was superior to the NQ or control treatment (no questions in text). No significant differences were found between the scores for the other treatment conditions (e.g., QR, questions - review and QNR, questions - no review).

Retention

The treatment conditions had a statistically significant effect on subjects' short term retention F(3,100) = 4.88, p < .05. Again, the Scheffe test indicated the mean (M = 6.00) for the QRW treatment was significantly higher (p < .05) than the mean (M = 5.01) for the NQ group with no other differences showing significance. No significant differences were found for the poor readers on short term retention F(3,58) = 0.374, p = .77.

While there were statistically significant differences between the long term retention scores of the subjects experiencing the various treatment conditions, the results of the Scheffe test failed to show significant differences. A similar pattern was obtained for both the scores of the good readers,
Interspersed Questions - 11

\[ F(3,39) = 2.40, p = .07, \] and the scores of the poor readers,
\[ F(3,58) = 1.63, p = .18, \] on the variable of long term retention.

**Discussion**

The findings will be discussed in conjunction with the research questions which have formed the basis for this investigation.

*Will interspersing questions in text enhance learning of actual content area material?*

The present investigation has used passages excerpted from actual content area textbooks. Since positive effects have been noted for interspersed question treatment QRW, it appears that interspersing questions in text, at least in the manner of treatment QRW, can enhance the learning of content area textbook material.

Most of the studies to date have employed contrived text developed to control for segment length and topical organization (e.g., Frase, 1968; Hudgins et al., 1979; Rickards, 1976). Yet, a need has existed to determine if the use of adjunct post-questions is effective with classroom material (Durkin, 1981; MacDonald-Ross, 1978; Rickards & Denner, 1978). Not only were the passages used in the present study derived from actual textbooks, but also the determination of question placement was based on the judgments of former junior high teachers. In previous studies (e.g., Hudgins et al., 1979; Rickards & DiVesta, 1974), question placement was determined by passage and segment length. Consequently, this study has provided evidence for the efficacy of interspersed questions using materials which were ecologically
representative of the classroom environment.

Will allowing students to review (look back in) the text facilitate comprehension?

The findings indicated that statistically significant differences existed between the control group and the questions-review-write group. No other significant differences were found between the remaining treatment group scores (questions-no review and questions-review). Therefore, it is not known if the effect resulted from forcing subjects' attention to the question through writing or if it was from the combined effect of writing with permission to review the text. An additional study could assess effects of a question-no review-write treatment in order to determine if review in combination with writing is the more effective treatment. Since no significant effects were noted for the question-review treatment, apparently just permitting students to review the text is not sufficient to enhance comprehension.

Will having students write their response, to ensure attention to the question, enhance comprehension?

In all instances where significant differences were noted between the scores of the subjects experiencing the various treatment conditions, the QRW treatment produced higher retention scores than did the control treatment. These findings were evidenced for all subjects (good and poor readers considered together) on two of the three dependent variables (short term retention and short term and long term retention combined). Treatment QRW was effective only for good readers and only on
short term retention and short term and long term retention combined. Since no significant differences were noted between the scores of subjects receiving either the question-no-review or the questions-review treatment, it appears that forcing attention to interspersed questions through writing is an effective method for improving comprehension. Most frequently in previous studies subjects were expected to make a covert mental response to the question (e.g., Rickards, 1976; Rickards & Hatcher, 1978; Rowls, 1974; Wilson, 1979). In the instances where an overt response was required, subjects were asked to identify a correct answer or fill in the blank with a word or phrase (e.g., Andre et al., 1980; Sagaria & DiVesta, 1978; Watts & Anderson, 1971). However, such procedures during reading are not common in a classroom situation.

Will good readers or poor readers receive the greater benefit from interspersing questions in text?

The results indicated that interspersing questions in text in the manner of treatment QRW is effective for good readers on short term retention and on short term and long term retention considered together. No significant differences appeared for poor readers experiencing the various treatment conditions.

A possible explanation for these findings with respect to good and poor readers concerns the interaction between the achievement level of the subjects and the readability level of the experimental materials. Since good readers were defined as scoring between 7.6 and 10.0 on the California Achievement Test and since scores on standardized tests often represent the
frustration reading level, it is likely that the seventh grade passages were at their instructional reading level. Therefore, with assistance in the form of questions interspersed in the text, good readers' comprehension was significantly enhanced.

Conversely, the poor readers, defined as scoring between 4.0 and 6.4 on the California Achievement Test, were experiencing material on their frustration reading level. As a result, even with the insertion of questions in text to reduce the amount of print encountered at one time, the readability level of the materials proved to be too difficult. These findings are at variance with the findings of Rickards and Hatcher (1978) who reported that meaningful learning postquestions, presented in the manner of treatment QNR, improved the comprehension of poor readers. They also reported that adjunct questions may be superfluous for good readers.

Numerous methodological differences exist between the Rickards and Hatcher study and the present investigation which may explain the disparity in findings between the two studies. A possible explanation centers around the definitions of and criteria used to differentiate between the good and poor readers under study. Rickards and Hatcher (1978), using 93 fifth graders as subjects, defined good readers as reading at or up to two years above grade level on the New Developmental Reading Test. Poor readers were defined as those subjects reading at least one year below grade level on the same test. Eliminated from consideration were those poor readers who scored at or below the third stanine on the word analysis section of the Stanford
Interspersed Questions

Diagnostic Reading Test. In the present study, seventh grade subjects reading between 6.5 and 7.5 grade levels were eliminated from consideration to maximize the differences between the two ability groups. Also eliminated from the data analysis were subjects whose reading grade equivalent scores fell below 4.0. It may be that the poor readers used in the Rickards and Hatcher study were superior in terms of reading ability to the poor readers used in this investigation which may account for the fact that they were able to comprehend a grade level passage which contained interspersed postquestions.

What effect will interspersing questions have on students' long term retention?

The post hoc analysis revealed no significant differences between the delayed retention scores of subjects experiencing the four treatment conditions. Consequently, it appears that interspersing questions in text does not significantly enhance students' long term retention scores. It may be that retaking all four tests during the same class period interfered with subjects' recall of the information. However, the data do not support the use of adjunct postquestions as aids for long term recall.

What effect will interspersed questions have on school age subjects' comprehension of textbook material?

In this study, positive results have been found for seventh grade subjects (defined as good readers) on the QRW treatment. Thus, it appears that interspersing questions in textbook material, in the manner of the QRW treatment, has a facilitative...
effect on school age subjects' comprehension. As indicated earlier, many of the studies on adjunct questions were conducted in a laboratory setting using college students who were either required to serve as subjects or who were paid to serve as subjects (e.g., Andre et al., 1980; Boyd, 1974; Rickards, 1976; Rothkopf, 1965). Other researchers (MacDonald-Ross, 1978; Rickards, 1970; Rickards & Denner, 1978; Durkin, 1981) have criticized this practice and have questioned the applicability of the findings to actual classroom settings. The present investigation has provided evidence to support the use of a QRW treatment with seventh grade subjects reading above grade level.

Implications and Conclusions

It has been suggested that empirical evidence is needed to validate many of the secondary reading strategies purported to facilitate comprehension (Patberg, 1979). The present investigation has attempted to provide evidence to either confirm or refute the use of interspersed questions in text. Since significant effects favoring a question-review-write (QRW) treatment has been demonstrated, it appears that questions inserted in text may be an effective instructional method, at least for some students.

The need to "slice the task," e.g., reduce the amount of print a student must deal with at a given time, by using questions inserted in text has been advocated by teacher educators (Readence & Moore, 1980). The present study has shown that this may be an effective instructional strategy for good readers (subjects reading above grade level) but not for poor readers (subjects reading below grade level) when the task also involved writing.
answers to the questions. Apparently, despite the use of interspersed questions as aids, poor readers still have difficulty comprehending content area textbook passages written on grade level.

Therefore, the findings of this investigation suggest that teachers must use adjunct aids in the form of interspersed questions very selectively. Due to the concept density of most content area textbook material even good readers seem to need. textbook modification techniques such as the QRW treatment. It cannot be assumed, however, that poor readers are receiving benefit from textbook material containing interspersed questions. Instead, poor readers may need alternative material written on an easier level in order to understand content area concepts.
References

Andre, T., Mueller, C., Womack, S., Smid, K., & Tuttle, M. Adjunct application questions facilitate later applications, or do they? *Journal of Educational Psychology*, 1980, 72, 533-543.


**TEST REFERENCE:**
