ILLINOIS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

PRODUCTION NOTE

University of Illinois at
Urbana-Champaign Library
Large-scale Digitization Project, 2007.

Technical Report No. 274

STRATEGIC DIFFICULTIES IN SUMMARIZING TEXTS

Peter N. Winograd
University of Kentucky

April 1983

Center for the Study of Reading

TECHNICAL REPORTS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
51 Gerty Drive
Champaign, Illinois 61820

BOLT BERANEK AND NEWMAN INC.
50 Moulton Street
Cambridge, Massachusetts 02238

The National Institute of Education
U.S. Department of Education
Washington, D.C. 20208

CENTER FOR THE STUDY OF READING

Technical Report No. 274

STRATEGIC DIFFICULTIES IN SUMMARIZING TEXTS

Peter N. Winograd
University of Kentucky

April 1983

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

The research reported herein was supported in part by the National Institute of Education under Contract No. HEW-NIE-C-400-76-0116 while the author was at the Center for the Study of Reading at the University of Illinois. The paper is part of the author's doctoral dissertation and considerable thanks are due to George McConkie, Ann Brown, Robert Tierney, Karl Koenke, Peter Johnston, and Jill LaZansky.

Abstract

This study examined the possibility that some eighth-graders' difficulties with the task of summarization may be linked to deficits in strategic skills. A systematic examination was made of the students' introspective awareness of the summarization task, ability to identify important elements in the text, and ability to transform the text into its gist.

Results indicated that most of the eighth graders were aware of the task demands of summarization. However, good and poor readers did differ in what they considered important, in what they included in their summaries, and in how they transformed the original text. Sensitivity to importance and efficient use of the transformations were significantly related to the ability to produce summaries. Sensitivity to importance was also significantly related to the ability to comprehend what had been read. The study suggests that when comprehension difficulties are encountered, teachers should assess the students' use of strategic skills and provide appropriate training if necessary.

Strategic Difficulties in Summarizing Texts

In the past few years researchers have begun to learn more about the higher-order comprehension problems that may contribute to some children's frustration with reading. The phrase "higher-order" refers to those problems that are not caused by inadequate decoding skills or problems in lexical access (Golinkoff, 1975-76). Recent research has emphasized a number of possible sources of such difficulties including those that may be related to differential language experiences, the lack of prior knowledge, or the lack of strategic skills (Collins & Haviland, 1979). It is the third possibility, a lack of strategic skills, which was the focus of this study.

The purpose of this study was to systematically examine the strategy differences between good and poor readers as they summarized what they had read. Three aspects of the use of strategies were of particular interest: (a) the awareness of the task demands involved in producing summaries, (b) the ability to identify important elements in the text, and (c) the ability to transform and reduce the full meaning of a text into its gist. These aspects provided the basis for three experimental hypotheses about why poor readers might have difficulty in producing adequate summaries:

(1) Poor readers are not aware that the purpose of summarizing is to convey the important ideas in a concise manner.

(2) Poor readers fail to identify the information which should, by adult standards, be included in a summary.

(3) Poor readers fail to use, or use ineffectively, those transformations used by their more fluent peers.

For each of these hypotheses there were three goals: (a) to identify and characterize possible strategy deficits evidenced by poor readers, (b) to examine the influence that such deficits might have on the ability to perform the task of summarizing, and (c) to examine the influence that such deficits might have on the ability to comprehend what has been read, in order to determine whether the use of the strategy extends beyond the task of summarizing.

For this study the ability to summarize is defined as the ability to "convey the main points concisely" (Webster's Third New International Dictionary, 1964). The main points are operationally defined as those ideas that were specifically identified as important in a rating or selection task by fluent adult readers or those ideas that were included most often in the summaries of fluent adult readers.

Summarization was selected as the experimental task for a number of reasons. First, the ability to get the gist of what one reads is of paramount importance and many students experience difficulty with the task (Kennedy, 1971). Second, sufficient research has focused on the strategies involved in summarization to provide the basis for a detailed task analysis (Brown & Day, Note 1; Kintsch & van Dijk, 1978). Third, although the relationship is unclear, some of the strategies used in summarization may also be used in comprehension so that information derived from studying summarization may inform us about comprehension processes in general (Kintsch & van Dijk, 1978; Johnson, 1978; Johnston, 1981).

The model proposed by Kintsch and van Dijk (1978) was selected as the conceptual framework for this study since it attempts to show how individual propositions in the text are transformed and condensed into the gist. The

essential components of this model are the reader's schema, the microstructure, the macrostructure, and the macro-rules. According to this model, readers progress through a text reducing and organizing its microstructure into a macrostructure through the application of a series of transformations known as macro-rules.

These macro-rules are not applied in a random manner, rather they are constrained by the reader's goals. The reader's goals determine which elements in the text are considered relevant and which are considered irrelevant. Elements may be assigned relevance according to two criteria (van Dijk, 1979). The first is textual relevance in which importance is defined in terms of what the author considers important. Such relevance is usually signaled through various cues in the text structure. The second is contextual relevance, where importance is based upon personal interests or background knowledge. Fluent readers are able to make use of both textual and contextual criteria so that importance is assigned to elements that are personally relevant and to elements the author intended to be relevant.

The research presented here is concerned with three interrelated aspects of the Kintsch and van Dijk model: (a) the appropriateness of the reader's goals, (b) the identification of those elements in the text which are most important and (c) the use of the macro-rules. These aspects were selected because evidence in the literature suggests that these three dimensions may be particularly difficult for poor readers.

First, studies indicate that some children's reading difficulties may be linked to their confusion about the appropriate goals of various tasks in reading (e.g., Baker & Brown, 1980; Downing, 1979). Most of the

supporting research has focused on the acquisition of early reading skills (e.g., Clay, 1969; Denny & Weintraub, 1963; Ehri, 1979; Johns, 1980; Reid, 1966). However, work by several researchers (e.g., Canney & Winograd, 1979; Meyers & Paris, 1978) has indicated that even older students may also be confused about some task demands of reading. Thus, some poor readers' difficulties in summarizing may similarly be due to their misconceptions about the task demands.

Second, studies have also demonstrated that good and poor readers differ in their sensitivity to importance (e.g., Dunn, Mathews, Bieger, 1979; Eamon, 1978-79; Meyer, Brandt, & Bluth, 1980; Smiley, Oakley, Worthen, Campione, & Brown, 1977). However, important questions about the nature and effects of these differences remain unanswered. Do poor readers differ in a systematic manner from good readers in what they consider important in texts? This is an important question because, to paraphrase Brown and Smiley (1977), judgments of importance could differ because the poor readers have a different conception of what is important or because they lack sensitivity to importance. The relationship between sensitivity to importance and reading comprehension performance is also unresolved. The issue here is whether or not sensitivity to importance accounts for individual differences in comprehension beyond that accounted for by lower-level decoding problems (Berger & Perfetti, 1977; Perfetti & Lesgold, 1977; Pichert, 1979).

Third, recent studies (Brown & Day, Note 1; Day, 1980; Tierney & Bridge, 1979) have demonstrated that good and poor readers differ in their ability to apply the various rules of summarization. Here too, questions

remain about the nature and effects of these differences. Which of the transformations are most difficult? How do differences in the ability to use the transformations relate to the ability to summarize what has been read or to comprehend what has been read?

In conclusion, although the literature suggests that good and poor readers differ along the dimensions of task awareness, sensitivity to importance, and use of the summarization rules, many questions remain unanswered. Moreover, no previous study has been made of all three dimensions using the same population. Thus, the task of summarizing and Kintsch and van Dijk's conceptual framework provide a unique opportunity to systematically study the strategic difficulties of poor readers.

Method

Subjects

The initial subject pool consisted of eighty eighth graders and forty adults. Poor readers were defined as those children who scored below the 50th percentile on the Reading Comprehension Subtest of the Stanford Achievement Test. The mean score on this test for the poor readers was 26.94 (*SD* = 9.84). Good readers were defined as those children who scored above the 59th percentile on the same test. The mean score on this test for the good readers was 66.28 (*SD* = 9.59). The adults were all associated with the University of Illinois, either as undergraduates, graduate students, or recent graduates at the doctoral level. Absences and incomplete data sets reduced the final numbers to thirty-six poor readers, thirty-nine good readers, and thirty-seven adults.

Materials

The first set of materials consisted of a series of multiple-choice and short answer interview questions designed to assess the subject's knowledge about the task of summarization. Questions #1 through #4 were asked before any of the other experimental tasks were administered.

(1) How often are you asked to summarize what you read in school?

(2) Has anyone ever taught you how to summarize what you read?

(3) If the answer to Question #2 is yes, what did they teach you to do?

(4) What does it mean to summarize an article?

The next three questions were asked after all the tasks were completed. Question #4 was repeated in order to assess any possible effects of the experiment itself.

(5) What does it mean to summarize an article?

(6) List at least 3 ways you decide which ideas from the article should be put in a summary.

(7) List at least 3 ways that your summaries are different from the original articles.

The next set of experimental materials consisted of eight expository passages adapted from trade books and elementary social studies, science, and reading texts (Education Development Center, 1970; Ginn Basic Reader, 1967; Holt Databank System, 1976a, 1976b; Noble & Noble, 1974; Wright, 1971). All of the passages were approximately equal in word length (M = 344, SD = 18.35) and they ranged from the upper third grade to the lower sixth grade in difficulty (Fry, 1977).

The eight passages were the basis for several of the other experimental materials. First, a five question multiple-choice test was developed for each passage. All five questions were designed to test comprehension of specific information in the passage. The stem and the correct answer were taken from information explicitly stated in the text.

Next, a word list was developed which contained 100 randomly selected unique words from the passages (excluding proper nouns). This list was designed to provide decoding rate and accuracy measures for each of the eighth-grade subjects.

The last set of materials was developed by taking each passage and simply listing its individual sentences. The sentences were then numbered consecutively and each was followed by a five point rating scale so that the subjects could rate how important the sentence was to the total passage.

Procedure

Each of the eighth graders was randomly assigned to one of four equal sized groups with the restriction of having equal numbers of good and poor readers in each group. The order of the passages was counterbalanced for each group.

The initial data on the eighth graders were collected in two stages over a three week period. Each child was involved in approximately eight sessions which usually ran about forty minutes in length. During the first stage, the children were individually tested to obtain measures of their decoding accuracy and speed. During this time they also answered the first set of written interview questions and completed the IPAT Culture Fair IQ test (Cattell & Cattell, 1960).

During the second stage, the eighth-grade students completed the majority of the experimental tasks. The general order of the tasks was the same for all of the eighth-grade subjects. First, the children read an article and recorded their reading time. Next they answered the five multiple choice questions without access to the article. When they had finished answering the questions, they wrote a sixty word summary of the article. They were given access to the article while they were writing the summary. Following a brief interpolated task, the subjects rated the relative importance of each sentence to the passage as a whole. During the final task, the subjects turned back to the numbered sentences and selected the five most important sentences in the whole article.

After the last step, the children were given a few moments to relax before going on to the next passage and repeating the process. Each child worked with a total of six of the eight passages and equal numbers of subjects read each of the eight passages. In addition, the summarization task and the rating and selection tasks were counterbalanced so that the children summarized and then identified importance for the first three passages and then reversed the order for the last three passages. When the final summary had been written, the children answered the remaining questions for the task awareness interview. All answers to the questions were written.

Six months later, the final data were collected on a subset of the good (N = 12) and poor (N = 12) readers. These children were asked to rate and select the important information from two passages they had encountered earlier. The purpose of this final task was to obtain a measure of the long-term reliability of the children's importance ratings.

The data on the adults were collected during a single two hour session. Each adult worked with all eight passages and all of the tasks were the same, except that they were not given the IPAT Culture Fair IQ test, the decoding measures, or the delayed importance rating task.

Scoring the Summary Protocols

The summary protocols were scored using a system which made it possible to identify which ideas from the original passage were included in the summary, as well as to record what transformations had been performed on those ideas. Fourteen kinds of transformations were identified in the summary protocols. These fourteen transformations were then collapsed into four broad categories: reproductions, combinations, run-on combinations, and inventions.

Reproductions refer to instances where subjects reproduced individual sentences in the original passage, usually through the use of paraphrase or word for word copying. Combinations were transformations where subjects had combined two or more sentences in the original passage into one sentence in the summary protocol. An example from one of the adult's protocols was:

> The river otter is very much like the sea otter but it is smaller, eats fish, frogs, and snakes, and travels overland.

Run-on combinations refer to instances where elements from several sentences in the original passage had been included in the protocol but in a less organized fashion than those transformations scored as combinations. This category was developed because it was necessary to distinguish between well-formed combinations and those produced by subjects who included words in a less organized yet linear fashion across sentences. An example from one of

the summaries produced by a poor reader will help clarify the distinction between run-on and regular combinations:

> Sea otters have low bodies short legs tails are thick and flat shape help them swim deep they eat from the sea eat fish frog snakes in water river otter smaller and not so heavy bend front legs and push back play slide down the river bank

Inventions refer to instances where subjects produced individual sentences which conveyed the meaning of a paragraph, several paragraphs, or even the whole passage. While these sentences did relate to the passage, it was very difficult to tie them to any specific elements in the surface structure of the original sentences. These inventions were often topic sentences which covered the information in a general way.

As a check on the reliability of the scoring system, a second rater checked 20% of the summary protocols produced by the three groups. The reliability figures for identifying which ideas in the original passages were included in the summaries were: .96 for the poor readers, .94 for the good readers, and .89 for the adults. The interrater reliability figures for agreement over which of the fourteen subcategories of transformations had been used was slightly less: .93 for the poor readers, .91 for the good readers, and .84 for the adults.

This scoring system differs from others currently in use (Brown & Day, Note 1; van Dijk, 1977) in a number of ways. The most important difference is that, in this system, transformations were categorized simply on the basis of how the original information was modified to produce the summary. No decisions were made, at this point, about the relative importance of the information that was modified as Brown and Day (Note 1) and van Dijk (1977)

have done. For example, Brown and Day (Note 1) use the term invention to refer to the production of a topic sentence for a paragraph which lacks one. In the present system, however, no determination was made about whether or not the transformation has resulted in a new topic sentence. The new sentence was scored either as a combination or as an invention depending on whether or not its elements could be linked to specific elements in the original passage.

Although the scoring system used in this study draws heavily upon the work of van Dijk (1977) and Brown and Day (Note 1), it is simpler to use on summaries produced from naturally occurring texts because raters have only to decide which transformation has occurred. Later, in a separate analysis, decisions can be made about the relative importance of the information which was modified.

Results

Analyses were performed on three sets of data. The first set consisted of the subjects' responses to the interview questions. The second set consisted of the subjects' ratings and selection of the important items in the eight passages. The third set consisted of the summaries themselves. Major findings are summarized below.

Fisher Z transformed correlation coefficients were used in all appropriate analyses. However, untransformed correlation coefficients are reported to ease communication and interpretation. Post hoc tests were Scheffe' with a significance level of .05. Tests of simple main-effects were performed for all significant interactions using the procedure recommended by Kirk (1968). In particular, the significance levels were

determined by dividing the overall alpha for the main-effect test evenly among the collection of simple main-effects tests.

Task Awareness

All subjects were asked a series of questions in an attempt to assess their awareness about the task of summarizing. The first two questions were multiple choice and were easily analyzed. For all the open-ended questions, the answers were analyzed by identifying responses given by more than one subject and then rank ordering those responses by the proportion of subjects who produced them. These rank-ordered responses were then further examined using the Goodman-Kruskal tau (Blalock, 1979).

Details regarding the subject's responses to individual questions can be found in Winograd (1982). However, 88% of the poor readers, 92% of the good readers, and 89% of the adults claimed that they are at least sometimes asked to produce summaries in school. In addition, 74% of the poor readers, 72% of the good readers, and 54% of the adults reported that they had been taught how to summarize.

Recall that the question, "What does it mean to summarize an article?", was asked twice in order to assess any possible effects of the experiment itself. The rather small changes in the proportional reduction of error between the first time and second time the question was asked indicated a relatively stable response pattern.

The analysis most pertinent to this study was conducted in the following manner. Theoretically, the most salient feature of a summary is that it contains the most important ideas in the original passage. Therefore, if a subject explicitly mentioned important points or main ideas

at least once in a response to any of the questions, he or she was assigned a score of 1 indicating an awareness of this aspect of summarizing. If the subject never explicitly made reference to importance, but made more vague responses about what the article was about, he or she was assigned a score of 2. This distinction was made so that it was possible to see if the more explicit reference to importance was an indication of a more developed metacognitive awareness of the task demands. If a subject never referred to importance or what the passage was about in response to any of the questions, he or she was assigned a score of 3. These scores are based on a nominal scale and were later used in multiple regression analyses to examine the possible effects of task awareness on various dependent measures.

Results indicated that 69% of the poor readers and 69% of the good readers were assigned response level scores of 1. Another 17% of the poor readers and 28% of the good readers were assigned response level scores of 2. The remaining children, 14% of the poor readers and 3% of the good readers, received response level scores of 3. Since so many of the students made explicit references to including important ideas as an aspect of summarizing or at least stated that a summary indicated what the article was about, it can be concluded that confusion about the task goal was not a major problem for these eighth graders.

It should be noted, however, that the task used to collect the data (written interview questions) can only measure task awareness on a crude scale. That is to say, although most of the children indicated that they knew what the general purpose of a summary was, a more sensitive measure may have detected subtle differences in the children's level of task awareness.

Sensitivity to Importance

The next set of analyses examined the subjects' ability to identify the important elements in the passages. Correlations were computed between the various experimental tasks at the within-individual level, between the individual and the aggregated group, and at the aggregated group level. Since the pattern of results was similar across all levels, most of the correlations discussed in this section are those computed between at the aggregated group level.

A preliminary comparison of the distributions of the poor readers' ratings with the distribution of the good readers' ratings indicated that both groups were using the full range of the five point rating scale and that the shapes of both distributions approximated the normal curve. Moreover, the within-subject variation for poor readers was about equal to that of the good readers. For example, after a six month lapse, the test-retest reliability coefficient for the poor readers equaled that of the good readers. The mean correlation between individuals' first and second ratings was .20 for both good and poor readers. Other evidence comes from point biserial correlations computed between each individual's importance ratings and importance selections. The mean point biserial correlations were .524 for the poor readers, .485 for the good readers, and .567 for the adults. Thus, given that the poor readers had made some judgments about the relative importance of the ideas in the passages and that these judgments were somewhat consistent, the issues surrounding poor readers' sensitivity to importance could be addressed.

The first major finding from this set of analyses was not unexpected. Good readers were better judges of importance than were poor readers when

that importance was defined in adult terms. Consider first the correlations computed between each group's mean importance ratings which are presented in Table 1. The correlation between the poor readers and the adults is .459; between the good readers and the adults it is .708. The correlations between the good and poor readers' mean importance ratings and the proportion of adults who selected the sentence as one of the most important reveals a similar pattern. Poor readers agree less with adults (.353) than do good readers (.625).

Insert Table 1 about here.

The second major finding was that, although poor readers' judgments of importance were not highly related to those of adults, poor readers did show some consistency in what they considered important. Some of the data to support this conclusion come from the correlations which were computed between the individual's ratings of importance and the mean ratings of importance of his or her peers. The mean correlations are .337, .370, and .616, for the poor readers, good readers, and adults, respectively. Consider also the correlation between each group's mean importance ratings and the proportion of that same group who selected the sentences as one of the five most important. These coefficients, which are presented in Table 1, are .770 for the poor readers, .808 for the good readers, and .838 for the adults.

The evidence argues strongly that poor readers in the eighth grade are not idiosyncratic in their judgments of importance. Given that poor readers have some consistent ideas about what is important in texts, the

next question is, "What kinds of information do poor readers consider important?".

In order to answer this question, a set of sentences which poor readers tended to select as important but which adults and good readers did not was identified using the Chi-square test. This analysis revealed significant differences ($p < .05$) between the groups for sixty-five (29%) of the two hundred and twenty-four sentences in the eight passages. These sixty-five sentences could be classified into four groups: (a) sixteen sentences selected by more fluent readers (good readers and adults) than by poor readers, (b) seventeen sentences selected by more children than adults, (c) seventeen sentences selected by more adults than children and, (d) fifteen sentences selected by more poor readers than fluent readers. These last fifteen sentences are the most pertinent to this study.

In examining the sentences chosen most often by poor readers, it is difficult to identify one overwhelming characteristic. However, the sentences do provide a sense of what appeals to the interests of the poor readers in this study. Consider, for example, the sentences selected by significantly more poor readers in the article entitled Cities in the 1800's.

Insert Table 2 about here.

It would appear that these sentences are full of highly visual detail. It seems that for poor readers, importance may have been based on factors which captured their interest. In contrast, significantly more fluent readers choose the sentences marked by a asterisk. Fluent readers seem to be defining importance more in terms of textual importance.

A third major finding was that the relationship between the information that poor readers judged to be important and the information that they included in their summaries was not very strong. That relationship was stronger for the good readers and stronger still for the adults. Some of the evidence for this conclusion comes from point-biserial coefficients obtained by correlating each individual's importance rating with a 1 or 0 depending on whether or not the ideas in that sentence were included in that individual's summary. Other evidence comes from phi coefficients obtained by correlating the dichotomous data for summary inclusion with the dichotomous data for importance selection (1 or 0 depending on whether or not the sentence was selected as one of the five most important). The point-biserial coefficients are .097, .161, and .357, for the poor readers, good readers, and adults respectively. The phi coefficients are .087, .155, and .331 for the poor readers, good readers, and adults respectively.

Still other evidence comes from the correlations presented in Table 1. The correlation between the poor readers' ratings and the poor readers' frequency of including information in a summary is .186. The parallel coefficients for the good readers and the adults were much higher, .598 and .750, respectively. Thus, while poor readers showed a level of consistency among themselves in what was regarded as important that approached that of the good readers, this level of judged importance had much less influence in determining what poor readers included in their summary.

The preceding results raise the question of why the correlation between importance and summary inclusion should be so low for poor readers. In order to answer this question, several analyses were performed. The

first examined the possibility that serial position effects had a strong influence on which ideas were included in the summaries. The sentences of each passage were divided into quartiles on the basis of their serial position with approximately equal numbers of sentences in each quartile. Next, the proportion of ideas included in the summaries was determined for each quartile. The results of this analysis are shown in Figure 1.

Insert Figure 1 about here.

The fluent adult readers display a characteristic curve which indicates that over a third (.34) of the ideas in their summaries came from the first quartile, then decreasing amounts from the second (.23) and third (.18) quartiles, and then an increase in the final quartile (.23). This curve is almost identical to the one obtained by Kintsch and Kosminsky (1977) in their work with summaries produced by college students. Good readers tended to approximate the pattern produced by adults, but with less tendency to include information from the end of the passage. Their proportions for the four quartiles were .35, .26, .20, and .17, respectively.

Contrast the adult and good readers' curves with that produced by the poor readers. Poor readers get almost half (.44) of their ideas from the first quartile, then steadily decreasing amounts from the second (.25) and third (.16) quartiles. Poor readers get .13 of their information from the final quartile whereas adults get .23 and good readers get .17.

Given these patterns for inclusion of information in the summaries, it is instructive to look at a parallel analysis which examines serial position effects on which ideas the subjects selected as being the most

important ones in the passages. These results are displayed in Figure 2. Again, note the characteristic curve for both the adults and good readers.

Insert Figure 2 about here.

The figures for the four quartiles for the adults are: .35, .22, .19, and .22. For the good readers, the figures for the four quartiles are: .29, .22, .21, and .24. The poor readers, however, selected about a quarter of their ideas from each of the four quartiles.

When the data displayed in Figures 1 and 2 were subjected to a 3 (Group) by 4 (Quartile) repeated measures analysis of variance, significant effects were found for both dependent variables. When the dependent variable was proportion of ideas included in the summary, the Quartile effect and the Group X Quartile effects were both significant, $F(3,327) = 84.59$, $p < .0001$, and $F(6,327) = 6.81$, $p < .0001$, respectively. When the dependent variable was proportion of important selections, the Quartile effect and the Group X Quartile effects were both highly significant, $F(3,327) = 14.75$, $p < .001$, and $F(6,327) = 3.97$, $p < .001$, respectively. Specific details on the tests of simple main-effects can be found in Winograd (1982).

These data suggest that poor readers are using two unrelated strategies; one for deciding what should be included in the summaries, and another for selecting which sentences are the most important. The patterns produced by the good readers and adults, however, suggests that they are using their sensitivity to importance to guide them in both the inclusion and selection tasks.

This interpretation of the data is supported by further analyses. The individual's mean importance rating for the ideas in each quarter of the

passage that he or she included in the summaries was computed. Next, each individual's mean importance rating for the ideas not included was computed for each quartile. These data are given in Table 3. Adults consistently

---

Insert Table 3 about here.

---

include in their summaries information that they rate as important. Good readers also included information that they rated as important in their summaries, and they deleted the information that they rated as less important. Poor readers, in contrast, choose higher rated information to include in their summaries only in the first three quartiles. As they progressed through the passage, the mean differences between the rating for inclusions and deletions became smaller and smaller until, in the last quartile, the information included in the summaries had a mean importance rating below that given to the information that was deleted.

These results are further evidence that fluent readers do rely on their sensitivity to importance in order to construct summaries. Less fluent readers, however, show much less consistency between what they include in a summary and what they rate as important. In addition, poor readers are much more prone to the adverse effects of serial position. That is to say, poor readers include less information from the latter part of the passage in their summaries and that which they do include is less important based on their own judgments.

Use of the Summarization Rules

The third set of analyses examined the possibility that poor readers failed to use, or used ineffectively, those summarization rules used by their more fluent peers.

Preliminary analyses revealed that the three groups did not differ in the number of words they used in their summaries, $F(2,109) = 1.2$, $p > .1$. They did, however, differ significantly in the proportion of ideas each deleted from the original passages, $F(2,109) = 10.39$, $p < .0005$. Post hoc tests revealed that the good and poor readers deleted significantly more of the original passage ideas than did the adults. The mean proportion of original passage ideas deleted by the poor readers, good readers, and adults were .78, .78, and .72, respectively. These data are an indication of the adults' superior ability to convey more ideas without using more words. Furthermore, the fact that the good and poor readers do not differ in the number of ideas deleted makes the comparison of which ideas are included more interesting.

The major finding for the third data set was that there were clear developmental trends in the use of each of the rules. This pattern is clearly displayed in Figure 3. The proportional means for the adults were .64, .11, .25, and .00 for combinations, inventions, reproductions and run-on combinations, respectively. For the good readers, the means were .38, .06, .39, and .17. For the poor readers they were .25, .04, .48, and .23. Note that increased reading skill led to fewer reproductions and run-on combinations and more combinations and inventions.

Insert Figure 3 about here.

This pattern was confirmed by statistical tests. A 3 (Group) by 4 (Rule Use) repeated measures analysis produced significant effects for Rule Use and the Group X Rule Use interaction, $F(3,327) = 60.99$, $p < .0001$, and $F(6,327) = 15.18$, $p < .0001$, respectively. Follow-up tests indicated that poor readers used significantly fewer combinations than did good readers, $F(2,436) = 35.5$, $p < .005$. This is important because combinations were the dominant transformations used by the adults. In addition, while the differences were not significant, developmental trends were evident in each of the other rules. The hypothesis that poor readers fail to use, or use less effectively, those rules used by more fluent readers has been supported.

The Link Between Strategy Use and Performance

The last set of analyses examined what influence differences in strategy use had on the ability to perform the task of summarizing and on the ability to comprehend what has been read. Hierarchical multiple regression (Cohen & Cohen, 1975) was used to construct and test a model using summarization and comprehension performance measures as the dependent variable. Since hierarchical regression assumes that variables are entered into the equation in a theory-governed manner, some explanation for the order of the variables is required.

The first variable entered into the equation was the IPAT Culture Fair IQ score. This measure of nonverbal IQ was included to account for the possibility that differences in performance were mainly due to differences in intelligence.

Speed and accuracy of decoding were the next variables to be entered into the equations. These two measures were obtained for each child when he or she read the word list aloud to one of the experimenters. Speed of decoding was the time in seconds it took the child to read all one-hundred words. Accuracy of decoding was simply the number of words missed. Decoding speed and accuracy were entered second and third because theoretically they are a necessary prerequisite for comprehension. In addition, decoding ability is often cited as the major source of individual differences in comprehension. The fact that IQ and decoding ability are entered before the variables of interest and are allowed to account for as much of the variance as possible provides for a stronger test of the hypotheses presented in this study.

The fourth and fifth variables in the model were orthogonal contrasts based upon the level of response given to the interview questions dealing with task awareness. The first contrast compared Response Level 1 with Response Level 2 and the second compared Response Levels 1 and 2 with Response Level 3.

The sixth variable entered into the equation was a measure of each child's sensitivity to importance. This measure was obtained by computing the correlation between that individual's ratings of importance and the mean adult ratings of importance.

The seventh variable in the model was a measure of effective rule use. This measure was obtained by computing the proportion of combinations and inventions out of the total transformations each individual used. Since fluent adults tended to use these two kinds of transformations most

often, it was thought that children who used a higher proportion of combinations and inventions were exhibiting a more mature pattern of rule use.

The remainder of the variables entered into the equations were two-way interactions between each of the main effects.

The order of the variables measuring task awareness, sensitivity to importance, and rule use was based on the comprehension model presented by Kintsch and van Dijk (1978, p. 372). They stated that:

> The reader's goals in reading control the application of the macro-operators. The formal representation of these goals is the schema. The schema determines which micropositions or generalizations of micropropositions are relevant and, thus, which parts of the text will form its gist.

For the purposes of the regression analysis it was assumed that a fluent reader comes to the task with the awareness that to summarize an article is to reduce it to its gist. Next, he or she identifies the ideas to be included, and by necessity those which can be deleted. Then as the reader identifies elements as relevant or irrelevant, the macro-rules are used to transform the passage into a summary.

Now that the rationale underlying the model has been discussed, the results can be presented. The first regression equation examined the relationship between strategy differences and the quality of the children's summaries. The measure selected to assess the quality of each child's summary was a summarization score obtained by computing the point-biserial correlation between the sentences the child included in his or her summary and the proportion of adults who also included the sentences in their summaries. Thus, those children who included sentences in their summaries

that were also included by a higher proportion of adults would have correlations approaching 1. Conversely, those children whose summaries were based on sentences not included by a large proportion of adults would have correlations approaching 0.

The results from the hierarchical regression equation using this correlation as the dependent variable are given in Table 4. The IPAT Culture Fair IQ test accounts for a significant proportion of the variance, 16.94%. Neither of the decoding measures reached significance, nor did either of the contrasts involving the level of task awareness. Sensitivity to importance did account for a sizeable proportion of the variance, 16.4%. Proportion of effective rule use also accounted for a significant proportion of the variance, 4.15%. None of the interactions reached significance.

Insert Table 4 about here.

The second regression equation examined the relationship between strategy differences and comprehension of the passages. These data are given in Table 5. The dependent variable is the average number of multiple choice questions the children answered correctly. The IPAT Culture Fair IQ test accounts for a significant proportion of the variance, 9.36%. Accuracy of decoding also accounted for a significant proportion of the variance, 4.98%. Neither contrast involving level of task awareness reached significance. However, the next variable, sensitivity to importance, accounted for 5.31% of the variance. Effective rule use accounted for an additional 3.55% of the variance, a proportion which approached significance.

Insert Table 5 about here.

Two interactions did reach significance. The IPAT IQ X Decoding Accuracy interaction accounted for 6.27% and the Decoding Speed X Response Level Contrast 2 accounted for 5.80% of the variance. An examination of these two interactions indicated that they were probably the results of ceiling effects since most of the children did very well on the multiple choice questions. This was certainly the case for the IQ X Decoding Accuracy interaction. However, the interaction between Response Level Contrast 2 and Speed of Decoding did produce an interesting finding. An examination of the data revealed one outlier who was almost two standard deviations above (slower than) that group's mean for speed of decoding and whose responses to the interview questions indicated that he thought that to summarize an article meant to memorize it. Apparently, this subject's view of the experimental task and his slow speed in decoding had interacted to such an extent that answering the comprehension questions at the end of the passages became very difficult. His mean comprehension score was 2 out of a possible 5 when the average for all subjects was 4.02.

The last regression equation examined the relationship between strategy differences and a dependent variable with no restriction on range--the children's score on the Reading Comprehension Subtest of the Stanford Achievement Test. These results are reported in Table 6. The IPAT Culture

---
Insert Table 6 about here.

Fair IQ test accounts for a large proportion of the variance, 20.08%. In contrast to the results of the previous regression analysis, speed of decoding, not accuracy of decoding, accounted for a significant proportion of the variance, 8.71%. Neither of the contrasts involving level of task awareness approached significance. However, sensitivity to importance did account for a significant proportion of the variance in the children's standardized reading scores, 5.4%, which was about the same proportion of the variance that sensitivity to importance accounted for in the regression analysis with passage comprehension as the dependent variable. Neither effective rule use nor any of the interactions achieved significance.

In considering the results from all three regression equations, the most striking result is that the ability to identify important elements in a text accounts for a significant proportion of the variance in all three dependent measures even after IQ and decoding ability are taken into account. This is strong evidence that higher-order comprehension difficulties may be linked to strategy deficits. This also indicates that the ability to identify important elements in a text is a strategic skill that underlies both comprehension and summarization.

Second, effective use of the rules seems to be a strategic skill that is more task specific. That is to say, the ability to reduce a passage into a summary through the use of the transformations identified in this study does not relate significantly to the ability to comprehend that passage. This conclusion is based on the fact that rule use only accounted for a significant proportion of the variance in the regression which used the summarization score as the dependent variable.

Third, for the most part, task awareness was not a significant problem in either producing summaries or understanding what was read. This is not surprising given that most of the children revealed that they were aware that one aspect of the task of summarizing was to include the main idea. The non-significance of the task awareness factor also emphasizes the importance of linking measures based on introspective data to measures of performance (Ryan, 1981). It otherwise would not have been difficult to assume that differences in the explicitness of the children's responses may have been indicative of underlying strategy differences.

Summary and Conclusions

Overall, the data did not support the first hypothesis that some poor readers' problems with summarization stem from confusion about the demands of the task. Results indicated that most eighth graders in this study knew that a summary should include the important ideas from a passage. The results also indicated that the level of explicitness in their definitions of summarizing was related little, if at all, to their performance in understanding or summarizing a passage. It is worth repeating, however, that a more sensitive measure than written interview questions might have resulted in a different conclusion.

The data did support the second hypothesis that some poor readers have difficulty in identifying the information that adults consider important. Good readers were more in agreement with adults in their conceptions of importance than were poor readers. Poor readers were, however, about as consistent in their judgments of importance as were good readers. Thus, the problem was not that poor readers were idiosyncratic

and lacked sensitivity to importance, but rather, that poor readers had different views about which ideas in a text were important.

Based on these results, one can speculate that fluent readers were able to identify importance based on judgments of contextual and textual constraints. In other words, even though they may have found some passage elements important because of their particular interests and backgrounds, they were also able to identify what the author considered important through the use of textual cues. In contrast, the less fluent readers were more likely to base their selections of important information based on contextual constraints only. The kinds of information they chose as important seemed to be those that were of high personal interest, not the kinds of information the author staged as more important in the passage. An important goal for future studies will be to identify more specifically the factors that influenced the poor readers' choices.

Another striking difference between good and poor readers was the degree of relationship between what the subjects identified as important and what they included in their summaries. One explanation for the low correlations for the poor readers is that the fluent readers were including ideas in their summaries based on what they perceived to be important throughout the passage, while the poor readers' choices of which ideas to include were adversely affected by serial position. It may be that, as the poor readers proceeded through the passage and the processing load became heavier, they became less adept at using their perceptions of importance to guide them in choosing which ideas to include in their summaries.

Another important finding concerning sensitivity to importances was that it accounted for significant proportions of the variance in the summarization scores and in both measures of comprehension ability even after differences in IQ and decoding ability were taken into account. These results provide reasonable evidence that higher-order comprehension difficulties may be linked to strategy deficits. These results also indicate that the ability to identify important elements in a passage is a strategic skill that underlies both summarization and comprehension.

Consider why sensitivity to importance is so vital. When fluent readers read, they are able to use textual cues and background knowledge to help identify important elements in the text. These important elements are then used to construct an internal representation of the author's message (Kintsch & van Dijk, 1978). Evidence presented earlier indicates that poor readers seem to have difficulty in using textual cues. Without such guidance it must be very difficult to construct an accurate, organized representation of what the author intended to communicate. Furthermore, given the importance of organization in memory, the lack of such a representation should make it less likely that information will be remembered initially, more likely that it will be forgotten in the interim, and less likely that it will be retrieved when needed (Meyer, 1977).

The data presented in this study also provided support for the third hypothesis that poor readers fail to use, or use ineffectively, those summarization rules used by more fluent readers. These data replicate and extend the results reported by Day (1980), Marshall and Glock (1978-79), and Tierney and Bridge (1979) which indicated that poor readers have

difficulty integrating individual propositions into larger units. In addition, the finding that effective rule use accounted for a significant proportion of the variance in the summarization scores but not the comprehension scores indicates that this strategic skill is more relevant to the task of summarizing than it is to the more global process of comprehension.

The fact that effective rule use was more related to the summarization scores than to the comprehension scores may help clarify the relationship between summarization and comprehension. One explanation is that the task of summarizing not only requires a reader to construct an internal representation of the author's message, as is required for comprehension but also requires that other, secondary decisions be made about the relative importance of the elements in the internal representation (Brown & Day, Note 1; Johnson, 1978). Moreover, it seems that these secondary operations require the active control of the reader to a much greater extent than do the comprehension processes which resulted in the internal representation initially. Poor readers run into difficulty with both stages of this task. Not only do they have difficulty in constructing an internal representation of the author's message, but they also have difficulty in the secondary operations required to produce a summary.

What educational implications can be drawn from this study? The first comes from the finding that sensitivity to importance accounts for significant proportions of the variation in the children's comprehension scores. Teachers may wish to assess children's sensitivity to importance when there is evidence of comprehension difficulties. The methods for this

assessment can be informal or formal. For example, simply asking the children about which information in the passage is the most important and why they think so may provide useful diagnostic information. A more formal method of evaluation might involve procedures similar to those used in this study where children could be asked to identify the most important sentences in a carefully selected test passage. The children's choices could then be compared with norms based on the choices of more fluent readers, or on results obtained from a theoretically based analysis of the text structure. In any case, teachers may find that some poor readers may need explicit training in higher-order comprehension strategies in addition to or instead of training on decoding skills. It is important to stress this point because, until recently, it has always been easier to focus on decoding problems since we have had a clearer understanding of what decoding skills need to be taught. This is not to imply that fluent decoding is not important; certainly, it is. However, additional skills, particularly those that deal with understanding large units of text, need to be taught. The findings presented in this study argue strongly that the ability to identify the important elements in a passage should be included as one of those skills. An important question for future research will be to find some effective instructional strategies for accomplishing this goal.

A second implication, closely related to the first, comes from the finding that poor readers did show some consistency in the kinds of sentences that they chose as important. This implies that while the sentences poor readers tended to select as important differed from those of adult and good readers, the selections were not necessarily idiosyncratic.

Rather, they seemed to be based on criteria other than those used by fluent readers. It may be necessary to lead poor readers to develop different and possibly more discriminating criteria for deciding what is important in a passage. Teachers may better facilitate this transition by showing greater sensitivity to the fact that what appears to be an incorrect selection given the perspective of the fluent reader, may in fact be a quite reasonable choice given the perspective of the less fluent or young reader.

The third implication is based on the data which indicate that the task of summarizing involves some strategies in addition to those required for comprehension. Therefore, when children have trouble summarizing what they have read, teachers should not automatically assume that the children are having difficulty in understanding what they have read. Although difficulties with the task of summarization may be symptomatic of comprehension problems, summarization difficulties are not necessarily confined to comprehension problems. It may be that some children's difficulties in summarization lie in the secondary operations used to condense and transform a passage into its gist. Thus, training these students in an attempt to improve their general comprehension abilities may not improve their performance on the specific task requirements of summarizing.

In conclusion, this study has provided information which should be useful in determining some of the sources of higher-order comprehension problems evidenced by many poor readers. The results reported here should also help to emphasize the notion that there is more to reading comprehension difficulties than inadequate decoding skills and that there is more to summarization than adequate comprehension.

Reference Note

1. Brown, A. L., & Day, J. D. Strategies for summarizing texts: The development of expertise. Unpublished manuscript, University of Illinois, 1980.

References

Baker, L., & Brown, A. L. Metacognitive skills in reading (Tech. Rep. No. 188). Urbana: University of Illinois, Center for the Study of Reading, November 1980.

Berger, N. S., & Perfetti, C. A. Reading skill and memory for spoken and written discourse. Journal of Reading Behavior, 1977, 9, 7-16.

Blalock, H. M., Jr. Social statistics. New York: McGraw-Hill, 1979.

Brown, A. L., & Smiley, S. S. Rating the importance of structural units of prose passages: A problem of metacognitive development. Child Development, 1977, 48, 1-8.

Canney, G., & Winograd, P. Schemata for reading and reading comprehension performance (Tech. Rep. No. 120). Urbana: University of Illinois, Center for the Study of Reading, April 1979. (ERIC Document Reproduction Service No. ED 169 521)

Cattell, R. B., & Cattell, A. K. S. Handbook for the culture fair intelligence test. Champaign, Ill.: Institute for Personality and Ability Testing, 1960.

Clay, M. M. Reading errors and self-correction behaviour. British Journal of Educational Psychology, 1969, 39, 47-56.

Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, N.J.: Erlbaum, 1975.

Collins, A., & Haviland, S. E. Children's reading problems (Reading Education Rep. No. 8). Urbana: University of Illinois, Center for the Study of Reading, June 1979. (ERIC Document Reproduction Service No. ED 172 188)

Day, J. D. Teaching summarization skills: A comparison of training methods. Unpublished doctoral dissertation, University of Illinois, Urbana, Illinois, 1980.

Denny, T., & Weintraub, S. Exploring first graders' concepts of reading. The Reading Teacher, 1963, 16, 363-365.

Downing, J. Reading and reasoning. New York: Springer-Verlag, 1979.

Dunn, B. R., Mathews, S. R., II, & Bieger, G. Individual differences in the recall of lower-level textual information (Tech. Rep. No. 150). Urbana: University of Illinois, Center for the Study of Reading, December 1979. (ERIC Document Reproduction Service No. ED 181 448)

Eamon, D. B. Selection and recall of topical information in prose by better and poorer readers. Reading Research Quarterly, 1978, 14, 244-257.

Education Development Center. Man: A course of study. Washington, D.C.: Curriculum Development Association, 1970.

Ehri, L. C. Linguistic insight: Threshold of reading acquisition. In T. G. Waller & G. E. MacKinnon (Eds.), Reading research: Advances in theory and practice. New York: Academic Press, 1979.

Fry, E. Fry's readability graph: Clarifications, validity, and extension to level 17. Journal of Reading, 1977, 21, 242-252.

Ginn Basic Reader. My do and learn book. Boston, Mass.: Ginn & Company, 1967.

Golinkoff, R. A comparison of reading comprehension processes in good and poor comprehenders. Reading Research Quarterly, 1975-76, 11, 623-659.

Holt Databank System. Inquiring about American history. New York: Holt, Rinehart & Winston, 1976a.

Holt Databank System. Inquiring about cultures. New York: Holt, Rinehart & Winston, 1976b.

Johns, J. L. First graders' concepts about print. Reading Research Quarterly, 1980, 15, 529-549.

Johnson, N. S. A structural analysis of the development of story recall and summarization. Unpublished doctoral dissertation, University of California, San Diego, California, 1978.

Johnston, P. Implications of basic research for the assessment of reading comprehension (Tech. Rep. No. 206). Urbana: University of Illinois, Center for the Study of Reading, May 1981.

Kennedy, E. C. Classroom approaches to remedial reading. Itasca, Ill.: Peacock Publishers, 1971.

Kintsch, W., & Kozminsky, E. Summarizing stories after reading and listening. Journal of Educational Psychology, 1977, 69, 491-499.

Kintsch, W., & van Dijk, T. A. Toward a model of discourse comprehension and production. Psychological Review, 1978, 85, 363-394.

Kirk, R. E. Experimental design: Procedure for the behavioral sciences. Belmont, Ca.: Wadsworth Publishing Co., 1968.

Marshall, N., & Glock, M. D. Comprehension of connected discourse: A study into the relationships between the structure of text and information recalled. Reading Research Quarterly, 1978-1979, 14, 10-56.

Meyer, B. J. F. The structure of prose: Effects on learning and memory and implications for educational practice. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N.J.: Erlbaum, 1977.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. Use of top-level structure in text: Key for reading comprehension of ninth-grade students. Reading Research Quarterly, 1980, 16, 72-103.

Meyers, M., & Paris, S. G. Children's metacognitive knowledge about reading. Journal of Educational Psychology, 1978, 70, 680-690.

Noble and Noble Basal Social Studies Series. Many Americans--One Nation. New York: Noble & Noble Publishers, 1974.

Perfetti, C. A., & Lesgold, A. M. Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, N.J.: Erlbaum, 1977.

Pichert, J. W. Sensitivity to what is important in prose (Tech. Rep. No. 149). Urbana: University of Illinois, Center for the Study of Reading, November 1979. (ERIC Document Reproduction Service No. ED 179 946)

Reid, J. F. Learning to think about reading. Educational Research, 1966, 9, 56-62.

Rumelhart, D. E. Understanding and summarizing brief stories. In D. LaBerge & J. Samuels (Eds.), Basic processes in reading: Perception and comprehension. Hillsdale, N.J.: Erlbaum, 1977.

Ryan, E. B. Identifying and remediating failures in reading comprehension: Toward an instructional approach for poor comprehenders. In T. G. Waller & G. E. MacKinnon (Eds.), Advances in reading research (Vol. 2). New York: Academic Press, 1981.

Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. Journal of Educational Psychology, 1977, 69, 381-387.

Tierney, R. J., & Bridge, C. A. The functions of inferences: An extended examination of discourse comprehension. In M. L. Kamil & A. J. Moe (Eds.), Twenty-eighth yearbook of the national reading conference. Clemson, S.C.: National Reading Conference, 1979.

van Dijk, T. A. Semantic macro-structures and knowledge frames in discourse comprehension. In M. A. Just & P. A. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, N.J.: Erlbaum, 1977.

van Dijk, T. A. Relevance assignment in discourse comprehension. Discourse Processes, 1979, 2, 113-126.

Webster's Third New International Dictionary of the English Language (Unabridged). Springfield, Mass.: Merriam, 1964.

Winograd, P. An examination of strategic difficulties in summarizing texts. Unpublished doctoral dissertation, University of Illinois, Urbana, January 1982.

Wright, R. H. Curious ways of common birds. New York: Lothrop, Lee & Shepard, 1971.

Table 1

Mean Correlations Between Group Mean Scores

for Various Tasks

Group	Group								
	Poor Readers			Good Readers			Adults		
	Mean Importance Rating	Importance Selection	Summary Inclusion	Mean Importance Rating	Importance Selection	Summary Inclusion	Mean Importance Rating	Importance Selection	Summary Inclusion
Poor Readers									
Mean Importance Rating		.770	.186	.660	.562	.368	.459	.353	.301
Importance Selection			.257	.557	.575	.380	.366	.326	.284
Summary Inclusion				.407	.443	.741	.461	.479	.587
Good Readers									
Mean Importance Rating					.808	.598	.708	.625	.579
Importance Selection						.591	.626	.659	.515
Summary Inclusion							.628	.603	.772
Adults									
Mean Importance Rating								.838	.750
Importance Selection									.729

Note. Correlations are based on 224 sentences.

Group importance ratings are the mean rating given to each sentence.

Group importance selection is the proportion of subjects who selected the sentence as one of the five most important.

Group summary inclusion is the proportion of subjects who included the sentence in a summary.

Table 2

Sentences Which Showed Significantly

Different Importance Selection Patterns

for Fluent and Less Fluent Readers

CITIES IN THE 1800'S

*In the last years of the 1800's, cities in the United States were growing faster than anyone had ever dreamed was possible. But as the cities grew, so did the problems.

One problem was slums, with crowded, dirty apartment buildings called tenements. In the slums, diseases spread quickly when people got sick. In crowded slums, people threw their garbage out the windows, where it grew into huge heaps in the streets and alleys. Insects and rats in the garbage caused more sickness in the slums.

With so many people in cities, garbage suddenly became a problem outside the slums. No one guessed that cities would ever have to find ways to collect the garbage. Why, even in New York, the biggest city in the country, garbage had always been eaten by pigs in the streets.

New buildings went up almost overnight. Many were poorly made and jammed close together. Most were made at least partly from wood. The danger of fire increased. Cities began to suffer from terrible fires that quickly burned down entire neighborhoods. Chicago had one of the worst fires. Most of the city was destroyed and hundreds of people were killed or hurt. If cities were going to be made safe, buildings had to be made better, and good fire departments were needed.

*Crime was another city problem. Oh, there had always been criminals. But like other people, criminals seemed to be especially attracted to the city. The only difference was that the criminals came for different reasons. Large numbers of people and businesses provided more targets for thieves. And great crowds made criminals hard to catch. Sometimes a gang

Table 2 (cont.)

would take over a neighborhood in the city and even the police were afraid to go there.

Nearly everyone could see that the new cities needed help. *But many people believed that it was not the job of the city government to solve the new problems like slums or garbage or crime.

Note. Underlined sentences were selected as important by significantly ($p < .05$) more poor readers than by good readers and adults.

Sentences with asterisks were selected as important by significantly ($p < .05$) more good readers and adults than by poor readers.

Table 3

Mean Ratings by Passage Quartile for Sentences Included in Summaries and for Sentences Deleted

Group		Passage Quartile			
		1	2	3	4
ADULTS	Inclusions	4.16 (.37)	3.95 (.44)	4.02 (.81)	4.06 (.40)
	Deletions	3.38 (.36)	3.33 (.38)	3.25 (.43)	3.34 (.42)
	Difference[a]	.79 (.30)	.63 (.29)	.77 (.73)	.72 (.29)
GOOD READERS	Inclusions	3.69 (.56)	3.56 (.56)	3.67 (.56)	3.56 (.73)
	Deletions	3.32 (.49)	3.30 (.46)	3.23 (.47)	3.36 (.45)
	Difference[a]	.37 (.49)	.26 (.31)	.44 (.50)	.20 (.70)
POOR READERS	Inclusions	3.34 (.60)	3.35 (.85)	3.38 (1.09)	3.12 (1.34)
	Deletions	3.14 (.58)	3.19 (.49)	3.28 (.49)	3.30 (.54)
	Difference[a]	.20 (.37)	.16 (.84)	.11 (.95)	-.18 (1.33)

Note. N of cases: adults = 37; good readers = 39; poor readers = 37.

Numbers in parenthesis are standard deviations.

[a]These are mean differences.

Table 4

Partitioning of Variance of Summarization Scores and Tests of Significance

Variable	F	Percentage of Variance
IPAT Culture Fair IQ	17.29***	16.94
Decoding Speed	<1	.05
Decoding Accuracy	1.50	1.47
Response Level Contrast 1	1.66	1.62
Response Level Contrast 2	<1	.04
Sensitivity to Importance	16.73***	16.40
Proportion of Effective Rule Use	4.23*	4.15
IPAT x Decoding Speed	<1	.08
IPAT x Decoding Accuracy	<1	.04
IPAT x Contrast 1	1.20	1.18
IPAT x Contrast 2	<1	.02
IPAT x Sensitivity to Importance	<1	--
IPAT x Proportion of Effective Rule Use	<1	.20
Decoding Speed x Decoding Accuracy	<1	.19
Decoding Speed x Contrast 1	<1	.02
Decoding Speed x Contrast 2	1.48	1.45
Decoding Speed x Sensitivity to Importance	<1	.79
Decoding Speed x Proportion of Effective Rule Use	<1	--
Decoding Accuracy x Contrast 1	<1	.85
Decoding Accuracy x Contrast 2	<1	.63
Decoding Accuracy x Sensitivity to Importance	<1	.94
Decoding Accuracy x Proportion of Effective Rule Use	<1	.97

Table 4 (cont.)

Variable	F	Percentage of Variance
Contrast 1 x Sensitivity to Importance	<1	.36
Contrast 1 x Proportion of Effective Rule Use	<1	--
Contrast 2 x Sensitivity to Importance	3.23	3.16
Contrast 2 x Proportion of Effective Rule Use	2.33	2.28
Sensitivity to Importance x Effective Rule Use	1.14	1.09

Note. $N = 74$, $R^2 = .55$.

All independent variables have one degree of freedom.

*$p < .05$.

**$p < .005$.

***$p < .0005$.

Table 5

Partitioning of Variance of Mean Passage Comprehension Scores and Tests of Significance

Variable	F	Percentage of Variance
IPAT Culture Fair IQ	9.09**	9.36
Decoding Speed	<1	.11
Decoding Accuracy	4.83*	4.98
Response Level Contrast 1	<1	.01
Response Level Contrast 2	2.23	2.29
Sensitivity to Importance	5.16*	5.31
Proportion of Effective Rule Use	3.44	3.55
IPAT x Decoding Speed	<1	.46
IPAT x Decoding Accuracy	6.09*	6.27
IPAT x Contrast 1	<1	.05
IPAT x Contrast 2	<1	.35
IPAT x Sensitivity to Importance	<1	.13
IPAT x Proportion of Effective Rule Use	<1	--
Decoding Speed x Decoding Accuracy	2.65	2.73
Decoding Speed x Contrast 1	2.79	2.87
Decoding Speed x Contrast 2	5.64*	5.80
Decoding Speed x Sensitivity to Importance	1.53	1.57
Decoding Speed x Proportion of Effective Rule Use	<1	--
Decoding Accuracy x Contrast 1	3.83	3.94
Decoding Accuracy x Contrast 2	<1	--
Decoding Accuracy x Sensitivity to Importance	<1	.01

Table 5 (cont.)

Variable	F	Percentage of Variance
Decoding Accuracy x Proportion of Effective Rule Use	<1	.76
Contrast 1 x Sensitivity to Importance	<1	.28
Contrast 1 x Proportion of Effective Rule Use	<1	.27
Contrast 2 x Sensitivity to Importance	<1	.93
Contrast 2 x Proportion of Effective Rule Use	<1	.27
Sensitivity to Importance x Effective Rule Use	<1	.33

Note. $N = 74$, $R^2 = .53$.

All independent variables have one degree of freedom.

*$p < .05$.

**$p < .005$.

Table 6

Partitioning of Variance of Stanford Achievement Reading Comprehension Subtest Scores and Tests of Significance

Variable	F	Percentage of Variance
IPAT Culture Fair IQ	21.01**	20.08
Decoding Speed	9.11**	8.71
Decoding Accuracy	3.71	3.55
Response Level Contrast 1	<1	.26
Response Level Contrast 2	2.23	2.13
Sensitivity to Importance	5.65*	5.40
Proportion of Effective Rule Use	1.83	1.75
IPAT x Decoding Speed	1.93	1.84
IPAT x Decoding Accuracy	<1	--
IPAT x Contrast 1	1.69	1.62
IPAT x Contrast 2	<1	.67
IPAT x Sensitivity to Importance	<1	.64
IPAT x Proportion of Effective Rule Use	1.60	1.53
Decoding Speed x Decoding Accuracy	<1	.19
Decoding Speed x Contrast 1	<1	.09
Decoding Speed x Contrast 2	<1	--
Decoding Speed x Sensitivity to Importance	2.23	2.13
Decoding Speed x Proportion of Effective Rule Use	<1	.14
Decoding Accuracy x Contrast 1	<1	.04
Decoding Accuracy x Contrast 2	<1	.13
Decoding Accuracy x Sensitivity to Importance	<1	.87
Decoding Accuracy x Proportion of Effective Rule Use	1.92	1.84

Table 6 (cont.)

Variable	F	Percentage of Variance
Contrast 1 x Sensitivity to Importance	<1	.31
Contrast 1 x Proportion of Effective Rule Use	<1	.08
Contrast 2 x Sensitivity to Importance	<1	.09
Contrast 2 x Proportion of Effective Rule Use	<1	.82
Sensitivity to Importance x Effective Rule Use	1.17	1.12

Note. $N = 74$, $R^2 = .56$.

All independent variables have one degree of freedom.

*$p < .05$.

**$p < .005$.

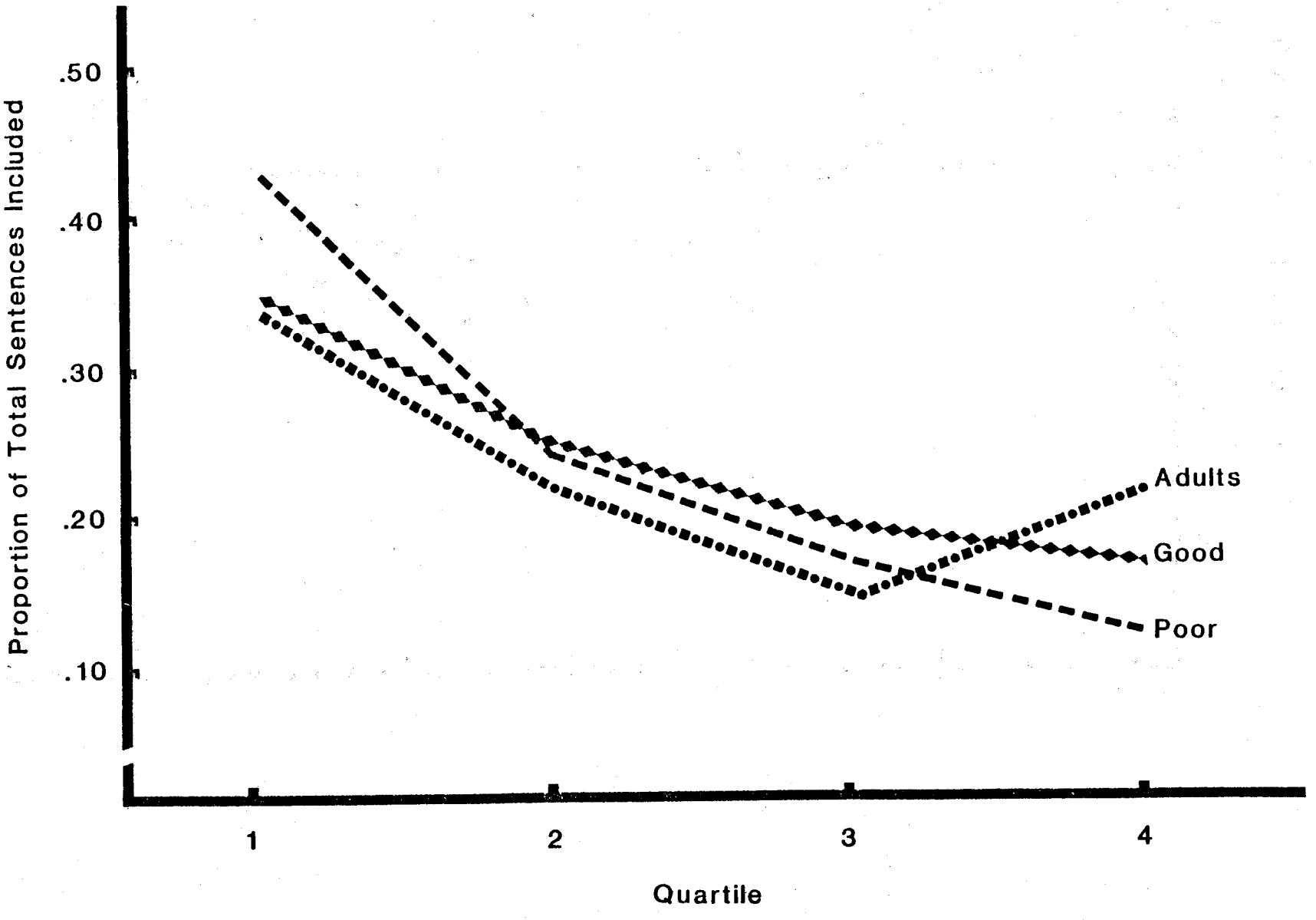
***$p < .0005$.

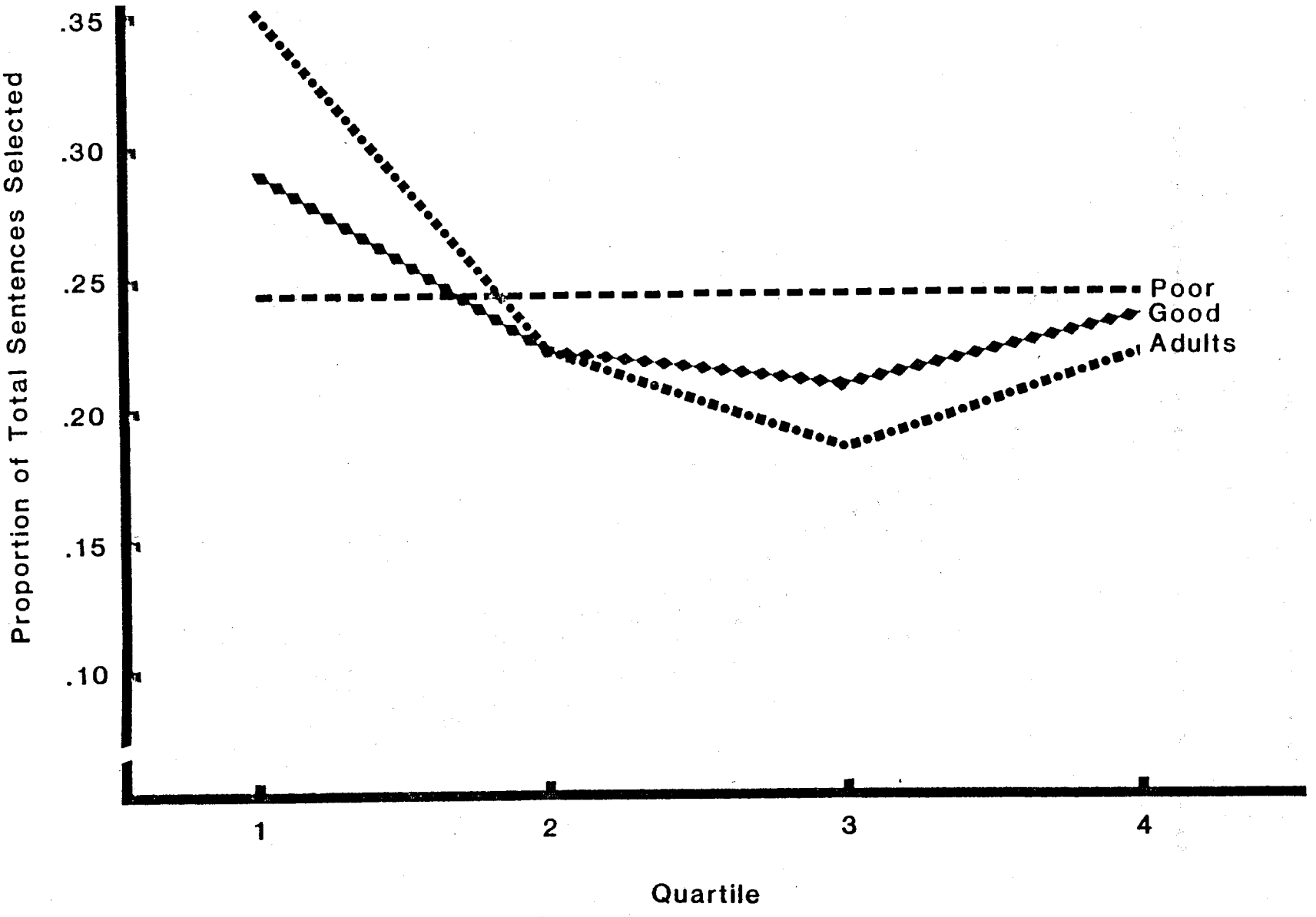
Figure Captions

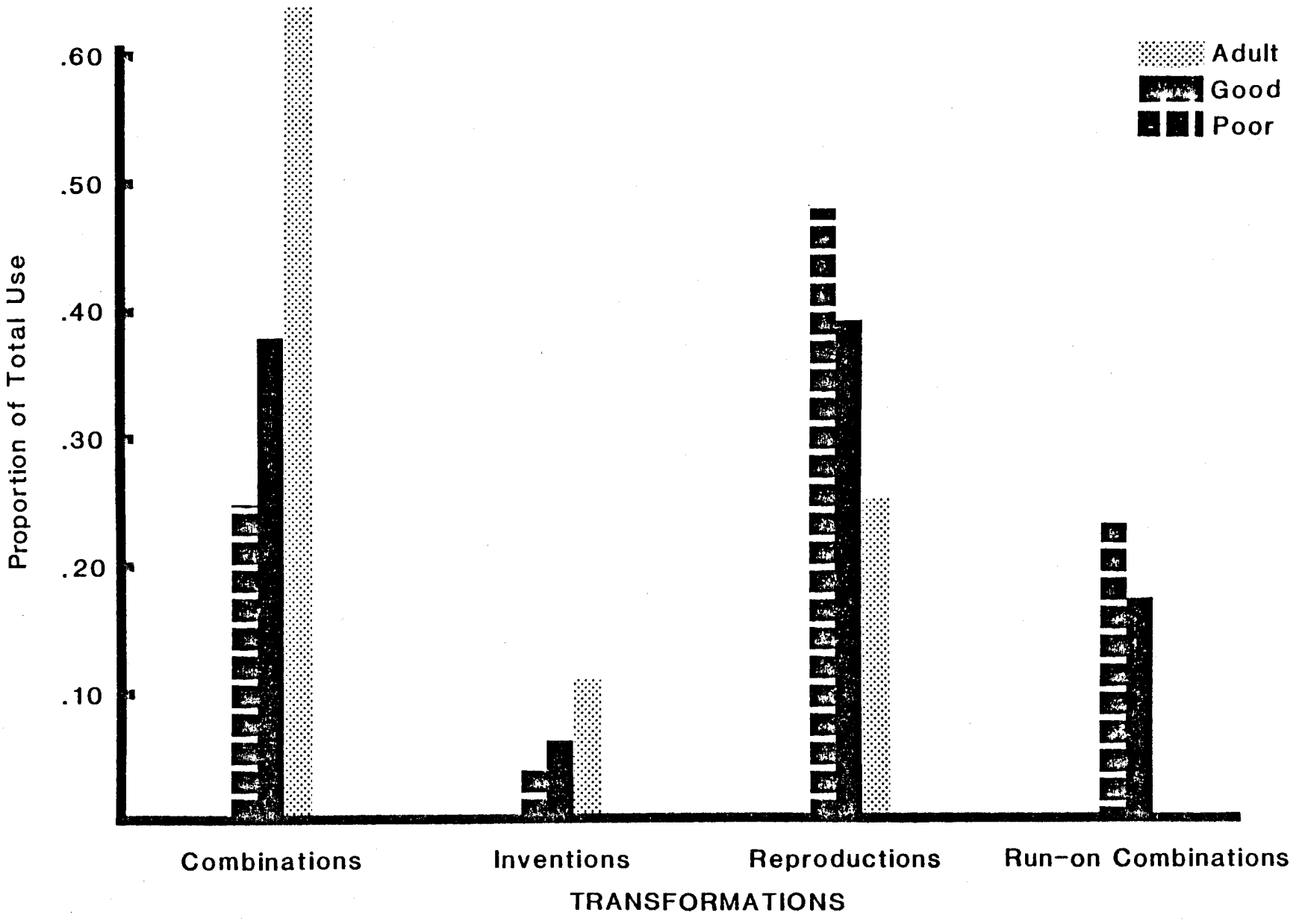
Figure 1. The proportion of sentences from each passage quartile which were included in the summaries.

Figure 2. The proportion of sentences from each passage quartile which was selected as the five most important.

Figure 3. Relative proportion of the use of the transformations.


Proportion of Total Sentences Included
.50
.40
.30
.20
.10
1
2
3
4
Quartile
Adults
Good
Poor


Proportion of Total Sentences Selected
.35
.30
.25
.20
.15
.10
1
2
3
4
Quartile
Poor
Good
Adults


Adult
Good
Poor
.60
.50
.40
.30
.20
.10
Proportion of Total Use
Combinations
Inventions
Reproductions
Run-on Combinations
TRANSFORMATIONS

This page is intentionally blank.