Macroprocesses and Microprocesses
in the Development of Summarization Skill

by

Eileen Kintsch

Communication Disorders and Speech Science

and

Institute of Cognitive Science

University of Colorado
Campus Box 345
Boulder, CO 80309-0345

ABSTRACT

The present study explored how students' mental representations of an expository text and the inferences they generate vary as a function of text difficulty and of differences in the task. Ninety-six students from Grades 4, 10, and College were asked to write summaries of an expository text and then to answer orally several probe questions about the content. Reading difficulty was systematically manipulated at the microstructure and macrostructure processing levels. The results supported the prediction of qualitative changes in the way the meaning is represented by different age groups in different text conditions. These are related to the amount and kinds of inferential processes on which the summaries were based. Interestingly, college students generalized the content more in summarizing texts with poor macrostructure than those that were well organized. The fact that more macropropositional statements occurred in responding to the probe questions than in the summaries could be explained in terms of the different retrieval conditions that prevailed. Some educational implications of these findings are discussed.

Macroprocesses   and   Microprocesses

in   the   Development   of   Summarization   Skill

Current views of reading comprehension assume an active process of
meaning construction at several interacting levels by which the reader derives
a personal interpretation of the textual content. This interpretation includes a
mental representation of the gist of the author's message. Thus, memory for the
essential information is what is typically retained by an experienced, adult
reader, especially following delay or if the passage is quite long.

The memory representation that is formed is the result of various
inferential operations which select certain kinds of information as important
and subsume details into more generalized statements. The van Dijk and W.
Kintsch (1983) model of discourse comprehension further assumes that the
reader constructs the generalized meaning, or *macrostructure*, on line during
reading, forming hypotheses about main points, or *macropropositions*, as soon as
there is enough information to hazard a guess. Subsequent research has
supported this assumption, at least for adult readers with normal skills who are
reading materials that are not excessively difficult (e.g., Guindon & W. Kintsch,
1984; Lorch, Lorch, & Mathews, 1985; Mross, 1988).

However, like other aspects of reading skill, efficient macroprocessing
strategies are the product of experience in dealing with textual materials of
various types. Not surprisingly, a number of recent studies have provided
evidence that school children, and even more advanced students, are not
efficient macroprocessors of school-type, expository texts. In fact, research on
the role of higher-level comprehension processes in reading suggests quite

strongly that this may form an important stumbling block to their learning (e.g., Bereiter, 1985: Brown, Bransford, Ferrara, & Campione, 1983; Perfetti, 1985).

Across a number of studies of children's text processing strategies a pattern has emerged that looks quite different from the gist-oriented strategies described in van Dijk and W. Kintsch (1983) and W. Kintsch and van Dijk (1978). Summaries that are largely generated by selection and deletion operations (e.g., Brown & Day, 1983; Winograd, 1984), failure to detect contradictions in meaning during reading (e.g., Markman, 1979; Scardamalia & Bereiter, 1984), associative writing that skips from topic to topic without an overall plan to guide it (e.g., McCutchen & Perfetti, 1982; Scardamalia & Paris, 1985), and incomplete recall protocols that ignore the hierarchical organization of text (e.g., Meyer, Brandt, & Bluth, 1980; Taylor & Samuels, 1983) are characteristic of immature readers and writers and those with inferior comprehension skill. Such findings, and others from a variety of school-related tasks, are apparently symptomatic of a general tendency among school-age students to process text in a linear, element-by-element fashion (e.g., Brown, Bransford, Ferrara, Campione,1983; Scardamalia & Bereiter, 1984). At least when dealing with expository prose materials, whose structures and content are relatively unfamiliar, such readers appear to concentrate most of their processing efforts on the meaning of individual sentences or sentence pairs, rather than on abstracting the global meaning (cf. also Englert, Stewart, & Hiebert, 1988; Di Vesta, Hayward, & Orlando, 1979; Taylor & Samuels, 1983).

In the present study the theoretical description of the comprehension process offered by van Dijk and W. Kintsch (1983) is used to characterize more fully the developmental differences that have been observed. The focus here, however, is not on strategy use per se, but rather on the crucial role of inferential processing in reading comprehension and on the relationship

between these meaning construction activities and the resulting memory representation. Scardamalia and Bereiter (1984; Bereiter & Scardamalia, 1987) have hypothesized that the linear processing strategy of younger, less skilled readers would result in a much shallower representation of the meaning than the multi-layered macrostructure of an expert. This study was designed to explore this issue under various text and task conditions.

As in many developmental studies of reading, a summarization task was chosen here to assess the underlying comprehension processes. It is assumed that summaries, like recall protocols, are a function of the way information is structured in memory. It should be noted, however, that summarization goes beyond recall in its demand to give explicit and coherent verbal expression to the conceptual representation of the meaning. By comparing both the inferred material and the structural properties of students' summaries with respect to the theoretical model of adult comprehension it is hoped that the relationship between level of processing and the resulting meaning representation can be clarified.

The summaries of an expository text written by students of three age groups were thus examined from two perspectives: The first is a fine-grained analysis of the text propositions and inferred content in the summaries; the second analysis examines text-based statements in both summaries and probe question responses in terms of their role in the hierarchy of macropropositions. The inference analysis provides a gauge of the constructive processing the material has undergone. The macrostructure analysis focuses on the mental representation of the text on which the summaries are based. Together, the two analyses are intended to reveal to what extent the meaning has been grasped at a macrolevel and the type of comprehension processes involved.

In addition to tracing the growth of macroprocessing ability with age, this study examined changes in the patterns of responses under different text conditions. Comprehension ease was systematically altered by manipulating the difficulty of the texts at the microstructure and the macrostructure level in order to gain a more detailed picture of the range of skills within each age group.

Finally, a probe question task was used to examine students' mental representations of text meaning under a different retrieval condition. Of interest here was to compare performance on the summaries with a task that supports the retrieval of the content via verbal cueing.

Method

Subjects

A total of 96 subjects participated in this study, 32 each from Grade 6, Grade 10, and college. The latter group were students from the University of Colorado; the school-age subjects were recruited from Boulder and Denver schools. Of the 6th grade subjects 17 were female and 15 male (mean age 11.6 years at time of testing). The high school and college groups both consisted of 20 female and 12 male students (mean age 15.6 and 18.9 years, respectively). All subjects were native speakers of English whose reading skills at each grade level generally reflected a high average within a normal range of abilities. Scores on national standardized tests or the passage comprehension subtest of the Woodcock Reading Mastery Tests (1973) were used to screen out subjects with substantial reading deficits.

Materials

Two passages of expository text were developed, each consisting of seven paragraphs (about 465 words). Both texts had the same underlying organization - a compare-contrast rhetorical structure - but they differed in content. Each text compared two developing countries in terms of their future potential. This

topic information was stated in the initial paragraph and at the conclusion of all versions of the texts.

There were four versions of each text. The basic version (Good Macro/Good Micro) used a tightly organized structure in which two countries were compared on the basis of three attributes: Geography, Economy, and Society or Culture. One text discussed Peru and Argentina, the other compared Indonesia and South Korea. The texts were written at an appropriate reading level for 6th grade students. This was determined informally on the basis of teacher evaluations and the 2% - 3% error rates of four average 6th grade readers who read the passage aloud.

The second version of the texts had a difficult macrostructure (Poor Macro/Good Micro), which was achieved by shifting from topic to topic instead of discussing various facts about each attribute in order. For example, a paragraph describing Indonesian agriculture, social problems, and trade was followed by a paragraph containing facts about Korean education, farming, and climate. Although the topic structure was disrupted in this way, local coherence relations were preserved. The version with a poor microstructure (Good Macro/Poor Micro) had more difficult and longer words, longer and more complex sentence patterns, and few explicit connectives signalling the relationships between sentences or phrases or between larger segments of text. Finally, the fourth version had both the macrostructure and the microstructure disrupted as described above (Poor Macro/Poor Micro). Examples of the materials are provided in the Appendix.

An important feature of the texts was the fact that the three major subtopics - the attributes on which the countries were being compared - were never explicitly mentioned and had to be inferred by the reader. (Though the scoring here is based on inferences, this would correspond to *topic invention* in

Brown and Day's, 1983, list of macrorules.) In addition, the content as a whole was stated at a very concrete level in order to maximize the opportunity for generalized statements in summarizing and in responding to the probe questions. Since the global topic was explicitly stated at the beginning and conclusion of each text, restatements or paraphrases of this information could be expected. (This would correspond to *topic selection* in Brown and Day.)

Differences in general knowledge undoubtedly existed across this range of ages, and although this could not be controlled in the present design, topic familiarity, was assessed informally. Furthermore, a neutral content domain was chosen, based on the assumption that most subjects, including the older ones, would have only rather vague knowledge about these foreign countries.

In sum, the experimental manipulations consisted of systematically interfering with comprehension processes at the microprocessing level and/or at the macroprocessing level in order to learn more about the skills that students of different ages bring to the summarization task: At what level are their processing efforts primarily directed under different text and task conditions? To what extent are they able to repair the problems exhibited by a particular text in their summaries?

Procedure

The subjects were individually tested in a single session lasting 30 - 50 minutes. Each subject was given one of the versions of either the Indonesia/Korea or the Peru/Argentina text to read and summarize in writing, with the conditions counterbalanced within each age group.

Brief oral instructions informed the subject about the nature of the task and provided them with a definition of the term *summary* (that it should state the important information in the text in a briefer form). Subjects were asked to read the passage at least two times and were allowed to underline the text, take notes,

or make a rough draft, if desired. Neither a time limit nor length constraints were imposed, and the text was available to refer to during writing. After finishing the summary, subjects filled out a brief questionnaire about their knowledge of the countries discussed in the text.

A set of four probe questions was then administered orally, and the responses were tape recorded for later transcription and analysis. The questions were designed to assess first in general terms a subject's understanding of the overall topic, and then more specifically what he or she considered to be the main ideas. The first question - "Can you tell me in a few words what this text is mainly about?" - could be answered by selecting one or more of the topic sentences that occurred in the text. The other questions probed for ideas, not all of which were stated in the text, especially the three implicit subtopics (Geography, Economy and Society/Culture). Thus, Questions 2 and 3 asked "On the basis of what you have just read, which country do you think will be better off in the future?" and "Why do you think so?" The last question was a specific probe for the implicit subtopics: "This article describes three important ways in which these countries are different. What do you think they are?"

Scoring

The summaries were analyzed in two ways. The first analysis was concerned with the number of actual text propositions vs. the inferred propositions in the summaries in order to provide an estimate of the meaning construction processes. The second analysis compared the the text-based information to the author's organization of the content to see to what extent the intended macrostructure was represented. Text was not a separate factor in either analysis; instead the results were collapsed across the two texts.

Text Propositions and Inferences. The summaries were first scored for the number of text propositions they contained. Using the propositional analysis

program developed by Turner (1987), each input text was decomposed into a list
of propositions which served as a template for scoring the textual information in
the summaries. A liberal, gist scoring criterion was employed (Turner & Greene,
1978). The nontext statements were then propositionalized and assigned to
inference categories.

Inferences form a continuum in terms of their closeness to the actual text,
as opposed to being extrapolations from the reader's own knowledge. However,
several categories of inferences can be defined according to the role they play
in the comprehension process. The following operational definitions are based
on van Dijk and W. Kintsch (1983).

(1) Generalizations are reductive inferences constructed from more
detailed statements in the text. They can be traced to the actual propositions they
subsume, except for global generalizations, which are inferences about the
overall meaning of the text. Generalizations reduce the number of text
propositions by at least one, though often by many more. For example, several
concrete statements about farming in Indonesia were often generalized as
Indonesians use primitive farming methods.

(2) Elaborations are inferences that are not directly implied by the text.
Instead they originate from the subject's own knowledge about the content of the
text or related information. For example, *Minerals are important to a country's
economy* cannot be traced to propositions in the Peru/Argentina text.

The other two inference categories were scored independently of the
elaborations and generalizations.

(3) Reorderings are inferences that re-arrange text content in a different
order than it occurred in the original text. Only between-paragraph, i.e., macro-
level, reorderings were considered here. These were not scored at the
propositional level; instead a sentence or phrase was counted as a reordering if it

required backtracking to an earlier paragraph to find its counterpart in the original text.

(4) Connectives are words that express bridging inferences, whose function is to provide coherence between the thoughts and ideas being expressed. There is much variability in the degree to which a writer or speaker makes these relationships explicit, but some inferential work is always left up to the comprehender. The number of explicit syntactic connectives between clauses, sentences, or paragraphs served as a measure of coherence processes. These were words or phrases, usually denoting a coordinating, causal, contrastive, or temporal relationship (e.g., *in addition, as a result, however, in contrast, since* etc.). A list of connectives, based on the linguistic description in van Dijk and Petöfi (1977), was used to guide scoring. The word *and* was not included in the analysis because it denotes a weak coordinating relationship and tends to be used as a filler or place keeper, especially by children.

To sum up, generalizations, reorderings, and elaborations are expected to play a particularly important role in summarizing the texts with poor macrostructure. Introducing connectives, which are an index of coherence-building processes, should be most evident in the summaries of poor microstructure texts from which they had been deleted.

Level in the Macrostructure. A macrostructure consists not only of generalizations of textual details, but also of propositions selected from the text on the basis of their importance to the overall meaning. Such propositions also function as macropropositions. Therefore it is important to examine both the amount of generalized information in a summary, and to see if the information included is macrorelevant. To what extent a summary or probe question response contains the important information from the text can be assessed by comparing it with the macrostructure that the author had in mind while writing the text.

Thus, in scoring both summaries and probe question responses, the level of importance was determined for each text-based statement by comparing it with this intended macrostructure.

The author's intended macrostructure for both the Indonesia/Korea and the Peru/Argentina texts consists of three levels of macropropositions and one level of details. The macrostructure of the Peru/Argentina text is provided in the Apppendix. The levels were scored as follows, with the number of possible statements at each level shown in brackets:

Level 1 (3 statements) consists of topic statements, either a label (e.g., *It's about Peru and Argentina*) or a more elaborated topic statement (e.g., that the two countries are compared, that there are factors that influence their development). This information could be found in the first and last paragraphs of the experimental texts.

Level 2 (3 statements) are the inferred subtopics, the three factors - Geography, Economy, Society - on which the countries are being compared. A statement at this level had to involve a comparison of the two countries.

Level 3 (15 statements) are other text-based macropropositions (e.g., Government, Education, Industry, Farming) that function as subheadings for groups of detailed statements. Some are inferred, others are mentioned in the text. These statements could refer to one country, or the two countries could be compared. For the latter credit was given for two macropropositions.

Level 4 (38 statements) consists of a representative, though not comprehensive list of concrete details from the text.

The summaries and responses to the probe questions were scored in the same fashion by matching the text-based statements they contained to counterparts at these four levels. Although the oral responses included many elaborations and sometimes spurious statements, only text-based statements could

be reliably assigned to a particular level. The score for both the summaries and the questions is based on the percentage of the total possible at each level.

The scoring was performed by the author. The list of text propositions for each text version was used as the basis for propositional scoring. A checklist was used to score the presence of connectives in the summaries, and reorderings were checked against the order of statements in the original text. Distinguishing between generalizations and elaborations and assigning statements to macrostructure levels involves subjective judgments. Therefore reliability in scoring these measures was assessed on 12% of the summaries by an independent rater who was blind to the experimental manipulations. The two raters agreed 84% and 88% of the time on the inference categories and the macrostructure levels, respectively. In both cases, most of the disputes were omissions or could be resolved in discussion.

Results and Discussion

A number of studies have pointed to qualitative differences in the way experienced and less experienced readers comprehend and use the information in expository materials. The present results agree with the conclusion drawn from this research that upper elementary school students are not very successful in their attempts to abstract a generalized meaning from school-type texts. Instead, their summaries show evidence of comprehension processes that operate largely on local level meanings. In the present study the developmental differences in summarization were characterized in terms of the inferential processes that produced the summary and the conceptual understanding of the text content that had been achieved. Specifically, an attempt was made to demonstrate how text features that promote or interfere with comprehension interact with age differences and task conditions in influencing the kind of

mental representation that is constructed during reading and the way the information is expressed in a summary.

The Analysis of Inferences. The number of text propositions, generalizations, elaborations, and reorderings and percent connectives in the summaries. were each subjected to a 3 (Age Group) x 2 (Macrostructure) x 2 (Microstructure) analysis of variance Because individual summaries varied considerably in length, the latter analysis is based on the percentage of connectives out of the total number of propositions rather than on frequencies. In addition, Newman-Keuls post hoc analyses were performed where significant main effects and interactions had been observed in order to specify the locus of the effect.

The Analysis of Macrostructure Levels. The percentage of text-based statements in the summaries and probe question responses occurring at each level of the macrostructure was entered into a 4 (Level) x 3 (Group) x 2 (Macrostructure) x 2 (Microstructure) multivariate analysis of variance. The percentage scores are based on the total possible at each level of the intended macrostructure. Level in the macrostructure was a within-subjects variable, and Group, Macrostructure, and Microstructure were between-subjects variables. These analyses were performed on arcsine transformations of the percentage scores, however untransformed means are reported below.

After testing for between subjects effects, a series of planned comparisons was performed to examine in more detail differences in the percentage of statements at each level as a function of age group and text condition. The three contrasts, which were chosen on the basis of theoretical interest, tested differences between the following pairs: (1) between topic-level macropropositions (Level 1) and the other two levels of macropropositions

(Levels 2 and 3), (2) between Level 2 and Level 3 macropropositions, and (3) between macropropositions (Levels 1, 2, and 3) and details (Level 4).

In the following section the results are presented first with respect to the effects due to age group that were observed on the two measures. The influence of text macrostructure and microstructure on these developmental differences is then discussed. Finally, the macrostructural representation of text content in the summaries is compared to that displayed by the probe question responses. A significance level of $p < .05$ was adopted for all analyses.

Age Effects in the Inference Analysis. The main effect of Age Group was significant on almost all of the measures investigated here, though not all differences between individual means were significant on post hoc analysis. An overview of the group means for the inference analysis is provided in Table 1. The number of text propositions included in the summaries formed the one exception which showed no systematic variation with age; nor were there any significant interactions. This is not surprising since subjects had access to the text while writing and hence their selection of textual information was not affected by memory constraints. Interesting differences in the summaries are captured, however, by the other measures.

----- See Table 1 -----

Group was a significant main effect in the number of generalizations contained in the summaries: $F (2,84) = 30.44$, MSe = 3644.89. As seen in Table 1, the younger students typically did not try to reduce the text information by generalizing it: The 6th graders produced significantly fewer generalizations than the older students according to post hoc tests. The high school students, in turn, generated somewhat fewer than the college students, though the difference between these means was not significant. A significant developmental effect was also observed in the amount of reordered material in

the summaries: F (2,84) = 10.54, MSe = 18.26. In general, the 6th graders were more likely to follow the order of information in the original text than the older students.

Significant group differences were also observed in the number of elaborations: F (2,84) = 6.61, MSe = 283.53 and the percentage of connectives included in the summaries: F (2,84) = 3.39, MSe = 4.05. As shown in Table 1, younger students did not elaborate what they had read as much as the older students. Nor did they supply as many bridging inferences, in the form of explicit connectives.

Thus, overall, the results point to a developmental trend in the amount of constructive processing that went into the summaries. In particular, the paucity of inferred information in the summaries of the elementary school students indicates that they were composed largely of information selected from the original text. This agrees with Brown and Day's (1983) finding that the predominate strategy used by upper elementary school students in summarizing expository texts is based on selection and deletion operations.

Age effects in the Levels Analysis. A significant main effect for Group occurred in the between subjects analysis of macrostructure levels: $F(2,84) =$ 29.29, MSe = 3.88. The elementary school students provided less text-based information than the older students. On the average, 27.51% of the text-based statements were included by the 6th graders, compared with 45.51%, and 50.37% for the 10th grade and college students, respectively. However, the main effect of Level and the Group by Level interaction were both significant in the within subjects analysis: $F(3,252)$ = 147.87, MSe = 15.53 for Level, and $F(6,252)$ = 10.93, MSe = 1.15 for the interaction of Level with Group.

The effect of Level is depicted graphically in Figure 1, which shows how the textual information in the summaries was distributed across the four levels of

the macrostructure, collapsed across the three age groups. Subjects generally provided a lot of topic information, but relatively few of the macropropositions at Level 2 and Level 3 of the macrostructure. Although a large number of detailed statements were included, the summaries contained proportionately more macropropositions than details overall. This effect was due to significant differences on two of the preplanned contrasts that compared the percentage of textual information at each macrostructure level. One was the contrast between Level 1 topic statements and the other macropropositions (Levels 2 and 3): $F(1, 84) = 275.56$. The other significant test contrasted the three macroproposition levels (Levels 1, 2, and 3) with the details (Level 4): $F(1,84) = 5.65$.

----- See Figure 1 -----

The significant Group x Level interaction showed up on two the three contrasts. For the contrast of macrorelevant information (Levels 1, 2, and 3 combined) with Level 4 details $F(2,84) = 14.28$. This pattern of results is shown in Figure 2. Here we see little variation in the amount of detailed information (Level 4) in the summaries across the three age groups, but the amount of macro-information (Levels 1, 2, and 3) increases substantially with age. An even sharper picture emerges when the scores are considered in terms of the proportion of the total text-based statements that occurred at the detail level. Dividing the Level 4 means of each age group by the summed means of all four levels reveals that 38% of the text material in 6th grade summaries was at the detail level, compared with 23% in 10th grade and 20% in college student summaries.

----- See Figure 2 -----

The results of this analysis thus confirm one of the observations made in the analysis of inferences. Despite differences in the unit of measurement - macropropositions or inferences - an increase occurs in the proportion of

generalized information with age, along with a corresponding decrease in the amount of detailed information.

The other significant interaction of Group with Level occurred in the contrast of Level 1 topic statements vs. Level 2 and Level 3 macropropositions: $F(2,84) = 12.08$. Converting these data into proportions reveals that 70% of all the text-based macropropositions provided by the 6th grade students were at the topic level, while there were rather few Level 2 or Level 3 macropropositions in their summaries. In contrast, macropropositions occurred at all three levels in the summaries of the older students. The predominance of topic information thus decreased to 62% among 10th graders and 58% among the college students.

In general, it appears that even the youngest subjects were able to select one or more topic statements to include in their summaries, but otherwise their summaries consisted largely of Level 4 details. It is worth noting here that of the three possible statements at Level 1, the younger students typically chose only the topic label (e.g., *It's about Peru and Argentina)*, while older students usually provided the elaborated topic statements as well. The results of the macrostructure analysis thus reveal a decrease not only in the importance of textual details, but also of topic information as the ability to form more specific macropropositions (Level 2 and Level 3) increases with age.

This result may be taken as evidence that the mental representations of the meaning constructed by younger vs. older, more experienced readers are indeed quite different, as Scardamalia and Bereiter (1984) have suggested. From the summaries of the youngest students one can infer a macrostructure composed primarily of a simple topic statement and an assortment of details from the text. Though a few other macropropositions may be constructed as well, these are generally not well integrated into the overall meaning. With age, however, there is an increasing tendency to differentiate several levels of importance in

the structure of the information. This conclusion is consistent with studies that have compared children and adults (Brown & Smiley, 1977) and good and poor readers (Winograd, 1984) on sentence rating tasks. Reading time data, collected by Lorch, Lorch, Gretter, and Horn (1987) indicate that school-age children do respond to the topic information in expository text in a manner comparable to adults. However, the present results suggest that the memory representation that they construct is not a well integrated, hierarchical structure, one that would provide an effective retrieval path in tasks where the information must be directly accessed.

The results thus show both qualitative and quantitative differences in the way the original text content is transformed in writing a summary by the three age groups. In particular, there is growth in the ability to reformulate the material in more general terms, which is in line with previous research on the development of summarization strategies (e.g., Brown & Day, 1983), but also in the ability to conceptualize the information at different levels of generality. Let us now consider how subjects of different ages were affected by the comprehension problems embodied in some versions of the texts.

Macrostructure Effects in the Inference Analysis. Significant Group x Macrostructure interactions occurred both for the number of reorderings $(F(2,84) = 7.80, MSe = 13.51$, and for the number generalizations $(F(2,84) = 3.50, MSe = 418.95)$ These results are shown in Figure 3.

----- see Figure 3 -----

As might be expected, there were more reorderings of the poorly organized texts $(\overline{\overline{X}} = 2.36)$ than of the well organized texts $(\overline{X} = 1.19)$. Post hoc testing revealed no significant differences between age groups in the Good Macro condition. However, in the Poor Macro condition the two older age groups

reordered the information in their summaries significantly more than the youngest subjects, who simply followed the text organization regardless.

Whereas the 6th graders produced somewhat fewer generalizations for texts with poor macrostructure than those with good macrostructure, these numbers were about the same for 10th graders. In contrast, the college students actually produced more generalizations in summarizing the Poor Macrostructure texts. The means for this age group were significantly different on post hoc tests. Though the poorly structured materials may have interfered somewhat with the comprehension of the youngest subject group, the differences here were not significant. However, these data do suggest that the oldest students tried to compensate for the the disorganized input by generalizing the information more than they did when the input was more orderly.

Macrostructure Effects in the Levels Analysis. Further support for this interpretation is provided by the levels analysis, which revealed a very similar interaction between age group and the organization of the text that was read. A significant interaction of Group x Macrostructure x Level occurred in the planned test which compared the percentage of macropropositions (Levels 1, 2 and 3) and details (Level 4): $F(2,84) = 3.12$. Figure 4 shows the expected increase with age in the amount of macro-relevant information for the Good Macro texts. The poorly organized texts appear to have differentially affected the three age groups. The amount of detailed information increased in the 6th grade summaries of these texts, while macropropositions decreased considerably. The opposite trend was evident in the college student summaries. The percentage of macro-relevant material remained about the same in the summaries written by 10th graders, though details also tended to drop out in the Poor Macro condition. The similarity of these results to those obtained with the generalization inferences strongly suggests that the oldest subject group had attempted to deal

with the poorly structured input by reformulating the information in more general terms.

----- See Figure 4 -----

Microstructure Effects in the Inference Analysis. Increasing text difficulty at the microstructure level caused a complex pattern of results. First of all, fewer connectives were produced overall for texts with the poor microstructure than for those with good microstructure: $\bar{X}$ = 6.61% vs. $\bar{X}$ = 8.74% ($F(1,84)$ = 9.11, MSe = 10.90). However, Figure 5 shows that the percentages were distributed differently across the three age groups and text conditions. Slightly fewer connectives were provided by the 6th graders in the Poor Micro condition, but there were relatively few connectives in their summaries in all text conditions. Connectives were present at about the same level in college student summaries of texts with good or poor microstructure. The largest effect, which was significant on post hoc tests, was observed among the high school students. Although they readily used connectives if the original text contained them, apparently they were not very successful at supplying them on their own, unlike the older students, who simply generated them as needed. In fact, the percentage of connectives in the high school summaries directly reflected their level of occurrence in the Good and Poor Micro texts: 11% vs. 5%, respectively.

----- See Figure 5 -----

A second effect of the texts with poor microstructure was evident in significant Group x Macrostructure x Microstructure interactions which were found on two inference measures: the number of elaborations ($F(2, 84)$ = 3.50, MSe = 150.07) and the number of reorderings ($F(2,84)$ = 7.76, MSe = 13.45). There is some similarity in these patterns, which may be seen in Figure 6. In comparison with the other age groups, the summaries of the college students contained a large number of elaborations - a difference which was significant on post hoc

tests - yet surprisingly few in the Poor Macro/Poor Micro condition. Though both of the older subject groups reordered the information in their summaries of poorly organized texts more than the 6th graders, the oldest group did so much less for texts in which both the macrostructure and the microstructure were poor, significantly less than the 10th graders in post hoc tests. A possible interpretation of this effect in terms of resource allocation will be considered in a later section.

----- See Figure 6 -----

Levels Analysis of Probe Question Responses. The probe question data pose an interesting contrast to the summaries in several respects. First of all, a more typical levels effect was observed on this task, which is shown in Figure 7. The main effect of Level was significant in the within subjects analysis: $F(3,252)$ = 130.34, MSe = 15.35, as were all three contrasts involving Level as a main effect. The values of F for each of the contrasts are the following: $F(1,84)$ = 109.01 for Level 1 vs. Level 2 and Level 3 macropropositions; $F(1,84)$ = 65.35 for the contrast of Level 2 vs. Level 3 macropropositions; and $F(1,84)$ = 485.29 for the contrast of Level 1, 2, and 3 vs. Level 4. Thus, the percentage of text-based statements decreased significantly at each of the four levels of the macrostructure in this task, which is a rather different pattern of responses than was observed in the summarization results, shown in Figure 1. Although Level 1 topic statements accounted for the largest amount of the information under both task conditions, many more of the inferred macropropositions (Level 2) were provided in responses to the probe questions than in the summaries ($\overline{X}$ = 49.69 vs. $\overline{X}$ = 21.50%), but somewhat fewer of the Level 3 macropropositions ($\overline{X}$ = 19.55% vs. $\overline{X}$ = 25.14%). Furthermore, the details from Level 4 were greatly de-emphasized in responding to the questions ($\overline{X}$ = 11.78%) as opposed to the summaries ($\overline{X}$ = 41.99%).

Age Group was a significant main effect in the between subjects analysis of the question responses as well: $F(2,84) = 32.00$, MSe $= 3.22$. Older subjects included more textual information in their responses than younger subjects at each level of the macrostructure, as was the case with their summaries. A second order Group x Level interaction was significant as well: $F(6,252) = 6.13$, MSe $= .72$. The locus of the interaction was in two of the contrasts: the test of macropropositions vs. details (Levels 1, 2, and 3 vs. Level 4): $F(2,84) = 25.18$ and the test of Level 1 vs. Level 2 and Level 3 macropropositions: $F(2,84) = 5.35$.

A comparison of the mean percentages of macropropositional and detailed information in the two tasks is presented in Table 2. As in the summarization task, the responses of college and high school students to the probe questions contained more macropropositions than those of the 6th grade students on the average. However, the age differences were much less extreme in the probe task: When directly questioned, the younger subjects were able to provide many more of the intermediate level macropropositions than before.

----- See Table 2 -----

Converting the percent scores into proportions of macrorelevant material in the responses again provides a clearer picture of what is going on (i.e., dividing the summed means of Levels 1, 2, and 3 by the summed means of all four levels). Though all subjects responded with more macropropositions in the probe condition than in the summaries, the magnitude of the increase was especially large among the youngest subjects. In fact, the proportion of macro-level information in the 6th graders' responses nearly reaches that of the older students: 91% macropropositions vs. 94% for 10th graders and 92% for college students. This pattern of results contrasts sharply with that observed in the summarization task in which the age differences were much greater: 6th grade

summaries consisted of 62% macropropositions, compared to 77% and 80% for the high school and college students, respectively.

Another interesting contrast with the summarization task is the lack of a significant Group x Macrostructure x Level interaction. Though age differences were evident in the amount of macrorelevant material produced by younger and older students in summarizing the texts with poor macrostructure, the quality of the input text apparently made little difference to subjects' ability to generate text-based macropropositions when directly probed for them.

Topic Familiarity

Developmental differences undoubtedly existed in students' knowledge about foreign countries (cf. Ohlhausen & Roller, 1988), but they are believed to have affected the present results in a rather general manner. Responses to questions about students' background knowledge of the countries discussed in the texts yielded more affirmative answers among the youngest students, many of whom had recently completed a social studies unit on some of these countries (13 elementary school students vs. 6 high school and 3 college students reported recent or specialized knowledge). Hence it is probably safe to assume that most subjects only had a rather vague, global knowledge of the topics they had read about.

Summary of Major Findings

In general, the results of both the inference and the macrostructure analyses support the prediction of a gradual increase with age in the amount of meaning construction processes that take place in summarizing, particularly in the degree to which the text content is re-interpreted at a macro-level. While younger children are able to relate details from the text to a global topic, older students differentiate more levels of importance in the information read. At the same time, difficulty at the macro- vs. microprocessing level affected the

summaries of the three age groups in rather different ways. The patterns of results indicate not only that the older students generated more inferences, but that they also seemed to control their processing more effectively. Namely, they generated the kinds of inferences that were important in this task, by generalizing and reordering the poorly structured texts and supplying missing connectives to improve micro-level coherence. The comprehension of the younger children was apparently disrupted by the texts with poor macrostructure, which resulted in even less generalized and more detail oriented summaries than otherwise. An example of a fairly typical 6th grade summary and a particularly good college student summary of a Poor Macro/Poor Micro text are provided in the Appendix to illustrate these two ends of the continuum.

The high school students formed an intermediate group, in that they had little trouble understanding these texts, and readily reordered the content to make a more coherent organization. However, they also tended to elaborate, rather than generalize the detailed content, and were less likely to repair the coherence problems of the texts with poor microstructure by making the meaning relationships explicit.

Finally, a strikingly different pattern of results was observed in the probed recall task than in the summaries. More macropropositional statements occurred in the question responses at all age levels, especially the implicit subtopics at Level 2 in the macrostructure. Although significant age differences were still present in the percentage scores, it is interesting that the youngest students emphasized macrorelevant, rather than detailed information in their responses almost to the same degree as the older students. Furthermore, the responses at all ages were unaffected by the quality of the text that was read.

Theoretical   Considerations

The theoretical model of discourse comprehension specified in van Dijk and W. Kintsch (1983) provides a framework for interpreting the rather complex pattern of results obtained here. In that theory discourse comprehension is described as an interactive cycle of text-driven and knowledge-based operations. Various kinds of inferences are involved, each playing a somewhat different role in the process of meaning generation. Bridging inferences, such as connectives, are necessary to form a coherent representation of the textual meaning. Elaborations, reorderings, and generalizations are considered optional, interpretive inferences. However, these kinds of inferences result in more extensive and deeper memory traces and are especially important when the content must be recalled or used in some way. Despite their integrative function, elaborative inferences are less desirable when the reader's goal is to recall the gist of the content or to write a polished summary. For these tasks the ability to generalize the detailed information and to recombine and reformulate the ideas are essential operations because they are directly involved in deriving a macrostructure.

That adult readers do indeed make these distinctions in inference types was shown in studies by Graesser and Clark (1985a). Though elaborative statements were quite common in the free recall protocols of their subjects, these tended to drop out of summaries of connected prose texts, while the number of generalizations increased. If we examine the overall frequencies across the three age groups in the present study we see that the 6th grade summaries contain about the same number of generalizations and elaborations   ($\bar{X}$ = 5.75 and $\bar{X}$ = 6.16, respectively). With age, however, the number of elaborations decreases as generalizations assume a greater role ($\bar{X}$ = 18.78 generalizations vs. $\bar{X}$ = 10.06 elaborations in 10th grade summaries; $\bar{X}$ = 26.91 generalizations vs. $\bar{X}$ = 12.00

elaborations among the college students). This would suggest that the older students, and especially the college students, are distinguishing the relative importance played by these two inference types in summarizing, whereas the 6th graders seem to treat summarization much like a recall task.

This distinction between the two inference types was not maintained by the 10th graders for the texts that were badly organized. Instead, these subjects resorted to the more primitive strategy of elaborating more in the Poor Macro/Poor Micro condition, as shown in Figure 6. In contrast, the college students dealt with the poor quality of the input by eliminating elaborations, thus reducing the overall length of their summaries. Since their summaries for the texts with poor macrostructure also contained more generalizations, the content thereby became somewhat more concise and more generalized.[1] The similar pattern of responses observed in number of reorderings could be due to the fact that there was less need to reorder the information when the content was restated at a more general level. However, this result could also be interpreted in terms of the cognitive demands of different inference types.

Capacity Limitations and Inferences

A number of studies of adults' processing of textual materials suggests that the occurrence of inferences is highly constrained by the capacity of the working memory buffer (e.g., van Dijk & Kintsch, W. 1983; Graesser & Clark, 1985b; Kintsch, W., & Vipond, 1979; Singer, 1988). At the same time, inferences differ in the amount of conscious effort they require. For a skilled reader coherence-building inferences are generated quite automatically (e.g., Graesser & Clark, 1985b; Just & Carpenter, 1980; Singer, 1988) whereas meaning construction inferences - those that generalize, reformulate, and extend the meaning on the basis of the reader's personal knowledge are more consciously controlled, and thus are costly in terms of processing space.

In the present study the number of inferences in all categories was found to increase with age, which suggests that resource limitations interfered with the younger students' ability to generate inferences. According to Perfetti (1985), an important aspect of developing skill in reading comprehension is the increased efficiency of lower-level verbal processes, which frees up memory capacity for more conscious, inferential processing. A number of studies have shown that younger students and those with poor comprehension skills are less likely to engage in spontaneous inferencing as they read (e.g., Bransford, Stein, & Vye, 1982; Oakhill, 1984; Paris, 1978; Weaver & Dickinson, 1982), which also supports this interpretation of the results obtained here. Even though decoding and accessing the lexical meaning were unproblematic for the 6th graders, these processes may have been less automatic for them than for the older students. Furthermore, the lack of connectives in the 6th grade summaries, and in the Poor Microstructure summaries of 10th grade students, suggests that middle-level coherence-building inferences may still have required some conscious effort. Hence insufficient processing resources may have affected the ability of these students to construct higher-level meanings or other inferences.

Increased processing load may also explain the decrease in elaborations, and possibly in reorderings as well, in college student summaries of the texts with poor macro- and poor microstructure. That is, in attempting to deal with the poorly written texts, the older students apparently concentrated their efforts on generating the appropriate kinds of inferences - generalizations and connectives. In contrast, the high school students reordered, but also elaborated more, while the youngest students simply produced fewer inferences. Generalizations, however, are more effortful than the other inference types. Hence the college students may have compensated for the extra processing costs by producing fewer elaborations and reorderings.

Skilled reading implies not only having available a repertoire of comprehension strategies, but also the ability to direct attention to different aspects of meaning construction as needed. What is clearly evident in the different patterns of inferences observed here is that elementary school students are limited both in their inferential processing of grade-level, expository materials, especially at the macro-level, and in their control of their own processing.

The Importance of Macrostructures in Comprehension

The surface-oriented strategies of the younger students in this study are reflected not only in the small numbers of constructive inferences in their summaries, but also in a shallow interpretation of the meaning: one consisting primarily of a global topic and assorted details. With increasing age and experience, students' summaries reveal a hierarchical organization that that more closely resembles the author's macrostructure.

Although there were significant age differences in the probed recall task, the patterns did not parallel the results of the summary analyses. Most notably, the proportion of macropropositional material increased substantially, especially in the responses of the younger subjects. Furthermore, the organization of the input text had no apparent influence on the responses. This result is not surprising when one considers the nature of the two tasks. Summarization, like recall, depends on having information well organized in memory. In fact, this is a major differentiator between such tasks and those involving cued recall and recognition memory, as Tulving (1983) has convincingly argued. There is considerable evidence that a macrostructure representation of the meaning plays an important role in facilitating the retrieval of information from memory. This structural framework apparently affects the quality of summaries as well. Even when the text is available to refer to during writing, as was the case

here, it is important to have a sense of what the important points are and how they are related in order to make judgments about what to include and how to generalize across the details The three age groups who participated in this experiment differed in their ability to generate a hierarchically organized structure for the information in the texts. However, the 6th graders apparently had the information available in memory and were able to respond appropriately when directly probed for it. For this task a well organized representation of the content is not as crucial because the questions themselves function as recall cues. Therefore, the fact that the younger students were more successful in generating macropropositions when answering the questions than in summarizing can largely be explained by the different retrieval demands of the two tasks.

These arguments are based on the theoretical assumption from the model of discourse comprehension, that summaries, like recall, reflect the way the information is organized in memory, which itself is the product of the operations that took place during comprehension. An alternative interpretation, it could be argued, is that the children's inadequate summaries reflect their faulty understanding of what a summary is. This, together with the added demands of the writing task, might explain why their performance was so much better in responding to the oral probe questions. However, several factors weigh against this interpretation. First, it should be noted that all subjects received a standard definition of what summarizing means and confirmed that they understood the task. Secondly, the developmental literature generally supports the notion that children's summaries reflect systematic changes in the strategic processing of the material, which are related to age and experience with informative types of materials. Several studies (e.g., Garner, 1985; Garner, Belcher, Winfield, & Smith, 1985; Winograd, 1984), which directly probed children's knowledge of

summarization task demands, indicate that by middle elementary school children have a good understanding of what summarizing means, and indeed, even young children are able to state the main events of a story when asked to summarize narrative materials (Brown, Day, & Jones, 1983; Johnson, 1973; Marshall, 1984; Taylor, 1986;). Apparently the major difficulty for younger students and poorer readers is with identifying the important ideas and especially, with trying to formulate main point statements of their own (cf. reviews by Hidi & Anderson, 1986; Garner, 1987). In addition to summarizing, the developmental lag in macro-level processing of expository text has shown up in a variety of other task situations, such as study strategies, note-taking, recall, and rating the importance of ideas (Brown et al., 1983; Winograd, 1984), comprehension monitoring and re-assembling scrambled text (Scardamalia & Bereiter, 1984), cloze comprehension (Di Vesta et. al., 1979), and essay writing (Bereiter & Scardamalia, 1987). However, in the later high school and college years, we see increasing evidence of the use of more efficient macroprocessing strategies, though many studies, including the present one, indicate that college students persistantly ignore the need to be concise in summarizing (e.g., Brown & Day, 1983; Garner & McCaleb, 1985).

If the contrasting results obtained here in the oral probed recall and summarizing conditions were due to oral vs. written performance measures rather than the underlying memory representation, one would expect the 6th graders to generate very brief summaries. Since the number of text propositions did not differ significantly across the three age groups, it appears that most of the children had no difficulty copying textual material into their summaries. Instead, the developmental differences could be characterized in terms of the amount and kind of inferential processing that took place. Scardamalia, Bereiter, and Goelman (1982) argue similarly that by later elementary school, text

production factors have little to do with differences in the quality of written compositions.

It is important to emphasize that the present research does not provide a direct test of the hypothesis that variable processing resources may be responsible for the developmental patterns observed here. However, the results are compatible with this interpretation and they also agree with the existing literature on inferential processing in reading comprehension tasks by adults and children. On the other hand, proficiency in processing verbal information is but one important factor that develops with age, for in addition, there is cognitive growth in the capacity to structure multiple pieces of information into larger, more complex and more abstract units (e.g., Fischer, 1980). This may also contribute to the large amount of concrete and detailed information in the summaries of younger students and the tendency to express the text information in the form of detail-subsuming macropropositions among the older students. According to Fischer and Kenny, (1986), one would also expect the complexity of a task to affect the level of performance similarly to what was observed here.

Still other factors undoubtedly enter into the ability to form macropropositions and other higher-level inferences as well. Indeed, the role of differences in general background knowledge, of experience with the structures of expository prose ,and of particular subject domains (e.g., Ohlhausen & Roller, 1988) have not been emphasized here. All of these factors contribute to the development of efficient information processing by making it easier for the comprehender to chunk incoming information into macropropositions.

Macropropositions contribute to the efficiency of processing verbal information by freeing up memory capacity. Macropropositions are superordinate concepts to which individual propositions can be related as they are encountered in the text. Macroprocessing strategies, though they are more

effortful, result in a well organized memory representation, which can serve as the framework for further operations. In contrast, when one attends primarily to local-level meanings, each incoming proposition is related only to the immediately preceding ones, and the resulting memory structure tends to be more fragmentary and less well integrated. The trick to being an efficient macroprocessor, however, depends both on being able to increase working memory space by chunking information into macropropositions and having sufficient processing resources to generate these higher-level inferences.

Educational Implications

Very different estimates of students' understanding of the materials they had read were obtained here, depending on whether they were questioned about the content or had to generate it themselves Though the discrepancy seemed surprisingly large, it quite agrees with well-understood principles of human memory. It further suggests a need to be sensitive to the different retrieval conditions evoked by the tasks used to assess student's knowledge, in research and educational settings alike.

Another intriguing finding of the present work was that the college students increased their level of constructive processing when comprehension was made more difficult, producing more generalized statements in their summaries than otherwise. In contrast, the youngest age group showed a tendency to fall back to lower-level strategies, producing fewer inference and a larger proportion of text details in their summaries of the disorganized texts. It would seem that for students with sufficiently developed text processing skills, the disorganized texts provided a challenge to put forth more constructive effort than the easy texts, whereas comprehension difficulty apparently had the opposite effect on students with less proficient skills. Some related findings have been reported recently by Mannes and W. Kintsch (1987) and McDonald (1988)

which lend support to the above conclusion. Both of these studies found that an advance organizer that was inconsistent with the structure or perspective of a text presented later benefitted college students' performance on questions that involved inferential reasoning, but interfered with the students' recall of the content. How are such results to be interpreted in light of endeavors to teach school age children higher-level macroprocessing strategies, such as deriving the main ideas in a text (e.g., Baumann, 1984; Hare & Borchardt, 1984)?

Although the results obtained here seem to support, once again, the argument that elementary and secondary students' macroprocessing would be enhanced by more readable school texts, there is another aspect to this issue which should also be considered. Namely, that educators need clearer definitions of the goals of instruction in terms of how a particular content is to be used. If the purpose is to assimilate a body of organized facts in order to reproduce them in that form on an exam, then a memory representation that reflects the structure of the learning material is certainly desirable, and materials and instructional methods that make it easier to derive the underlying text macrostructure are clearly called for. However, if students are expected to abstract general meanings from individual facts, to apply the content to novel situations or to perceive relationships with other pieces of knowledge, to judge the validity of an argument or to defend a position, then they would benefit from having to impose their own structure on what they are studying. This extra cognitive effort would probably result in a memory representation that is more richly integrated into the personal knowledge base, and hence more adaptable to different situations.

There is a dual challenge here for educators, who, on the one hand must help less experienced students acquire the processing tools that will enable them to abstract generalized meanings and the important information from a text,

even in nonoptimal reading situations. At the same time, an instructor must find ways to support the students' attempts to construct their own interpretations of that meaning. Our research can contribute to this endeavor by concerning itself with the conditions that enable constructive learning to take place for students who are at various levels in the development of expertise.

References

Baumann, J. F. (1984). The effectiveness of a direct instructional paradigm for teaching main idea comprehension. *Reading Research Quarterly, 20,* 93-115.

Bereiter, C. (1985). Children need more complete reading strategies. In J. Osborn, P. T. Wilson, and R. C. Anderson (Eds.), *Reading education: Foundations for a literate America.* Lexington, MA: D. C. Heath.

Bereiter, C., and Scardamalia, M. (1987). *The psychology of written composition.* Hillsdale, NJ: Erlbaum.

Bransford, J. D., Stein, B. S., and Vye, N J. (1982). Helping students how to learn from written text. In M H. Singer (Ed.), *Competent reader - disabled reader. Research and applications.* Hillsdale, NJ: Erlbaum.

Brown, A. L., Bransford, J. D., Ferrara, R. A., and Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell and E. M. Markman (Eds.), *Handbook of child psychology,* Vol. 3 (4th ed.). New York: John Wiley and Sons.

Brown, A. L., and Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior,* 22, 1-14.

Brown, A. L., Day, J. D., and Jones, R. S. (1983). The development of plans for summarizing texts. *Child Development, 54,* 968-979.

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

van Dijk, T. A., and Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

van Dijk, T. A., and Petöfi, J. S. (Eds.) (1977). *Grammars and descriptions*. Berlin/New York: Walter de Gruyter.

Di Vesta, F. J., Hayward, K. G., and Orlando, V. P. (1979). Developmental trends in monitoring text for comprehension. *Child Development, 50*, 97-105.

Englert, C. S., Stewart, S. R., and Hiebert, E. H. (1988). Young writers' use of text structure in expository text generation. *Journal of Educational Psychology, 80*, 143-151.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review, 87*, 477-531.

Fischer, K. W., and Kenny, S. L. (1986). Environmental conditions for discontinuities in the development of abstractions. In R. S. Mines and K. S. Kitchener (Eds.), *Adult cognitive development: Methods and models*. New York: Praeger.

Garner, R. (1985). Text summarization deficiencies among older students: Awareness or production ability? *American Educational Research Journal, 22*(4), 549-560.

Garner, R. (1987). *Metacognition and reading comprehension*. New York: Academic Press.

Garner, R., and McCaleb, J. L. (1985). Effects of text manipulations on quality of written summaries. *Contemporary Educational Psychology, 10,*, 139-149.

Garner, R., Belcher, V., Winfield, E., and Smith, T. (1985). Multiple measures of text summarization proficiency: What can fifth graders do? *Research in the Teaching of English, 19*, 140-153.

Graesser, A. C., and Clark, L. F. (1985a). Structures and procedures of implied knowledge. In R. O. Freedle (Ed.), *Advances in discourse processes* (Vol. 17). Norwood, NJ: Ablex.

Graesser, A. C., and Clark, L. F. (1985b). The generation of knowledge based

inferences during narrative comprehension. In G. Rickheit and H.

Strohner (Eds.), *Inferences in text processing*. .Amsterdam: North Holland.

Guindon, R., and Kintsch, W. (1984). Priming macropropositions: Evidence for the

primacy of macropropositions in the memory for text. *Journal of Verbal

Learning and Verbal Behavior, 23*, 508-518.

Hare, V. C., and Borchardt, K. M. (1984). Direct instruction of summarization

skills. *Reading Research Quarterly, 21*, 2-78.

Hidi, S., and Anderson, V. (1986). Producing written summaries: Task demands,

cognitive operations, and implications for instruction. *Review of

Educational Research, 56*(4), 473-493.

Johnson, N. (1983). What do you do if you can't tell the whole story? The

development of summarization skill. In K. E. Nelson (Ed.), *Children's

language* (Vol. 5). Hillsdale, N. J.: Erlbaum.

Just, M. A., and Carpenter, P. A. (1980). A theory of reading: From eye fixations to

comprehension. *Psychological Review, 87*, 329-354.

Kintsch, W., and van Dijk, T. A. (1978). Toward a model of text comprehension and

production. *Psychological Review, 85*, 363-394.

Kintsch, W., and Vipond, D. (1979). Reading comprehension in educational

practice and psychological theory. In L. G. Nilsson (Ed.), *Perspectives on

memory research*. Hillsdale, NJ: Erlbaum.

Lorch, E. P., Lorch, R. F., Jr., Gretter, M. L., and Horn, D. G. (1987). On-line

processing of topic structure by children and adults. *Journal of

Experimental Child Psychology, 43*, 81-95.

Lorch, R. F., Jr., Lorch, E. P., and Mathews, P. D. (1985). On-line processing of the

topic structure of a text. *Journal of Memory and Language, 24*, 350-362.

Mannes, S., and Kintsch, W. (1987). Knowledge organization and text organization. *Cognition and Instruction, 4,* 91-115.

Marshall, N. (1984). *The effects of discourse type on summary statements by children of different grade levels.* Paper presented at the meeting of the American Educational Research Association, New Orleans.

McCutchen, D., and Perfetti, C. A. (1982). Coherence and connectedness in the development of discourse production. *Text, 2,* 112-139.

McDonald, D. (1987). *Drawing inferences from expository text.* Unpublished Ph.D. dissertation, New Mexico State University, Las Cruces, NM.

Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development, 50,* 643-655.

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

Mross, E. F. (1988). *Macroprocessing in expository text comprehension.* Unpublished Ph.D. dissertation, University of Colorado.

Oakhill, J. V. (1984). Constructive processes in skilled and less skilled comprehenders' memory for sentences. *British Journal of Psychology, 6,* 335-354.

Ohlhausen, M. M., and Roller, C. M. (1988). The operation of text structure and content schemata in isolation and in interaction. *Reading Research Quarterly, 23,* 70-88.

Paris, S. (1978). The development of inference and transformation as memory operations. In P. A. Ornstein (Ed.), *Memory development in children.* Hillsdale, NJ: Erlbaum.

Perfetti, C. A. (1985). *Reading ability.* New York: Oxford University Press.

Scardamalia, M., and Bereiter, C. (1984). Development of strategies in text processing. In H. Mandl, N. L. Stein, and T. Trabasso (Eds.), *Learning and comprehension of text*. Hillsdale, NJ: Erlbaum.

Scardamalia, M., Bereiter, C., and Goelman, H. (1982). The role of production factors in writing ability. In M. Nystrand (Ed.), *What writers know: The language, process, and structure of written discourse*. New York: Academic Press.

Scardamalia, M., and Paris, P. (1985). The function of explicit discourse knowledge in the development of text representations and composing strategies. *Cognition and Instruction, 2*, 1-39.

Singer, M. (1988). Inferences in reading. In M. Daneman, T. Waller, and G. MacKinnon (Eds.), Reading research: *Advances in theory and practice* (Vol. 6). New York: Academic Press.

Taylor, B. M., and Samuels, S. J. (1983) Children's use of text structure in recall of expository material. *American Educational Research Journal, 20*, 517-528.

Taylor, K. K. (1986). Summary writing by young children. *Reading Research Quarterly, 21*(2), 193-207.

Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.

Turner, A. A. (1987). *The propositional analysis system* (Tech. Rep. No. 87-2). Boulder: University of Colorado, Institute of Cognitive Science.

Turner, A. A., and Greene, E. (1978). *The construction and use of a propositional text base*. JSAS Catalogue of selected documents in psychology, MS 1713.

Weaver, P. A., and Dickinson, D. K. (1982). Scratching below the surface structure: Exploring the usefulness of story grammar. *Discourse Processes, 5*, 225-243.

Winograd, P. (1984). Strategic differences in summarizing texts. *Reading Research Quarterly, 19,* 404-425.

Woodcock, R. W. (1973). *Woodcock reading mastery tests revised.* Circle Pines, MN: American Guidance Service.

Appendix

Examples of Texts

Peru and Argentina, Good Macro/Good Micro

Many factors influence whether a developing country can look forward to a prosperous future, or whether it will be forever doomed to backwardness and poverty. A comparison between Peru and Argentina illustrates this point.

The lofty, snowcapped peaks of the Andes Mountains cover most of Peru. Even the vallyes are so high that the air is thin and cold. Where the mountains reach down to the Pacific Ocean, it never rains, and the air is so dry that even wooden tools and cloth made hundreds of years ago are perfectly preserved in the sand.

In contrast, the heartland of Argentina consists of a huge, grassy plain, called the Pampas, where cowboys herd cattle on ranches as large as those in Texas. Where rivers flow into the Atlantic, the capitol city sprawls, with stately buildings and broad avenues, almost like in Paris.

Wheat fields cover much of the Pampas, whose soil is so rich that it never needs fertilizing. However, Argentina, like America, has trouble finding buyers for all the wheat and beef it produces. The country has almost no coal, iron or other minerals. Thus it cannot manufacture many things that people need, such as cars, machinery, and clothing, so these items must be imported from other countries.

Peru is very different. There are many small farmers who grow just enough corn to feed their families. Other people work on large plantations where coffee, cotton, and sugar cane is produced for export to other countries. The mountains have rich deposits of copper, silver, and lead, and the government has developed some very profitable mines. The Indians who live in

the high mountain valleys raise sheep and llamas. The women weave beautiful sweaters and blankets from the wool of the llamas, just as their ancestors did. These weavings eventually find their way to the fashionable boutiques of Europe and the U.S. where they bring high prices.

Most Peruvians are Indians, descendants of the once proud and mighty Incas. When the Spanish destroyed the Inca empire, the Indians became the poor and oppressed people in their own land. A small group of white plantation owners has ruled the country ever since. They brought with them Christianity and the Spanish language, but the great Indian masses remained outside the Spanish culture. Most of them do not even speak Spanish, and few can read or write.

In Argentina, on the other hand, there are very few Indians today. The streets in the big cities are crowded with busy, energetic people, and in the outdoor cafes lively discussions can be heard. In recent years, the newspapers and magazines have been free to publish everything, for the people were finally able to establish a free, democratic government.

Thus Peru and Argentina may develop very differently in the future.

Peru and Argentina, Good Macro/Poor Micro

Many factors influence whether a developing country can look forward to a prosperous future, or whether it will be forever doomed to backwardness and poverty. A comparison between Peru and Argentina illustrates this point.

The towering, snow-encrusted peaks of the Andes Mountains extend over most of Peru. At this altitude, even the valleys are cold, and the air is thin. Where the mountains thrust downward to the Pacific, there is no moisture in the air, and wood implements, and even cloth, made hundreds of years ago, are perfectly preserved in the sand.

The Pampas is the Argentinian heartland, a huge, grassy plain where cowboys herd cattle on vast, Texas-style ranches. On a river estuary on the Atlantic Ocean the capitol city sprawls, with stately buildings and broad avenues that rival those of Paris.

Wheat fields cover much of the Pampas. The alluvial soil is so rich that fertilization is unnecessary. Argentina, like America, produces more wheat and beef than it can consume or even market abroad. The country is not blessed with deposits of coal, iron, or other minerals. Large-scale manufacturing of goods needed by the population is impossible. They must depend on other countries for imports of automobiles, machinery, and clothing.

Many Peruvians are small farmers whose families manage to subsist on what they grow. Others are employed on large plantations, where coffee, cotton, and sugar cane is produced for export. The mountains are riddled with copper, silver, and lead deposits. Extracting the ore has proven to be a profitable venture for the government. Sheep and llamas are raised by Indians who live on the high mountian plateaus. Luscious sweaters and blankets are created by the women from alpaca and wool on hand looms, just as their ancestors did. These garments eventually make their way to the fashionable boutiques of Europe and the U.S. where they command high prices.

The majority of Peruvians are Indians, descended from the once proud and mighty Incas. The Inca empire was destroyed by the Spanish conquerors, and the Indians became the impoverished and oppressed people in their own land. The country has been ruled by a small nucleus of white plantation owners ever since. Along with Christianity they tried to impose the Spanish language on the natives. The Indian masses have remained alientated from the Spanish culture. Most are unable to speak Spanish, and few have learned to read or write.

Only a small Indian minority lives in Argentina today. The bustling city streets are crowded with busy, energetic people. In the outdoor cafes lining the boulevards, animated discussions can be heard. In recent years it has been possible for newspapers and magazines to publish everything. The people finally succeeded in establishing a free, democratic government.

Thus Peru and Argentina may develop very differently in the future.

Peru and Argentina. Poor Macro/Good Micro

Many factors influence whether a developing country can look forward to a prosperous future, or whether it will be forever doomed to backwardness and poverty. A comparison between Peru and Argentina illustrates this point.

Most Peruvians are Indians, descendants of the once proud and might Incas. Most of them do not even speak Spanish, and few can read or write. However, the Andes Mountains have rich deposits of copper, silver, and lead, and the government has developed some very profitable mines.

The capitol city of Argentina sprawls where rivers flow into the Atlantic, with stately buildings and broad avenues, almost like in Paris. Unlike Peru, very few Indians live there today. Argentina, like America, has trouble finding buyers for all the wheat and beef it produces. The streets in the big cities are crowded with busy, energetic people, and in the outdoor cafes lively discussions can be heard.

The lofty, snowcapped peaks of the Andes Mountains cover most of Peru. Even the valleys are so high that the air is thin and cold. The Indians are the poor and oppressed people in their own land. Since the Spanish destroyed the Inca empire, a small group of white plantation owners has ruled the country.

On the other hand, wheat fields cover much of the Argentinian heartland, called the Pampas. The soil is so rich that it never needs fertilizing. The Pampas consists of a huge, grassy plain, where cowboys herd cattle on ranches as large

as those in Texas. In recent years, the newspapers and magazines have been free to publish everything, for the people were finally able to establish a free, democratic government. The country has almost no coal, iron or other minerals. Thus is cannot manufacture many things that people need, such as cars, machinery, and clothing, so these items must be imported from other countries.

In contrast, the Indians who live in the high mountain valleys of Peru raise sheep and llamas. The Spaniards brought with them Christianity and the Spanish language, but the great Indian masses remained outside the Spanish culture. The women weave beautiful sweaters and blankets from the wool of the llamas, just as their ancestors did. These weavings eventually find their way to the fashionable boutiques of Europe and the U.S. where they bring high prices.

Where the Andes Mountains reach down to the Pacific Ocean, it never rains, and the air is so dry that even wooden tools and cloth made hundreds of years ago are perfectly preserved in the sand. There are many small farmers who grow just enough corn to feed their families. In addition, many Peruvians work on large plantations where coffee, cotton, and sugar cane is produced for export to other countries.

Thus Peru and Argentina may develop very differently in the future.

Peru and Argentina, Poor Macro/Poor Micro

Many factors influence whether a developing country can look forward to a prosperous future, or whether it will be forever doomed to backwardness and poverty. A comparison between Peru and Argentina illustrates this point.

The majority of Peruvians are Indians, descended from the once proud and mighty Incas. most are unable to speak Spanish, and few have learned to read or write. The Andes Mountains are riddled with copper, silver, and lead deposits. Extracting the ore has proven to be a profitable venture for the government.

The capitol city of Argentina sprawls on a river estuary on the Atlantic Ocean, with stately buildings and broad avenues that rival those of Paris. Only a very small Indian minority lives there today. Argentina, like America, produces more wheat and beef than it can consume or even market abroad. The bustling city streets are crowded with busy, energetic people. In the outdoor cafes animated discussions can be heard.

The towering, snow-encrusted peaks of the Andes Mountains extend over most of Peru. At this altitude, even the valleys are cold and the air is thin. The Indians are the impoverished and oppressed people in their own land. Since the Spanish conquerors destroyed the Inca empire, a small nucleus of white plantation owners has ruled the country.

Wheat fields cover much of the Argentinian heartland, called the Pampas. The alluvial soil is so rich that fertilization is unnecessary. The Pampas consists of a huge grassy plain, where cowboys herd cattle on vast, Texas-style ranches. In recent years it has been possible for newspapers and magazines to publish everything. The people finally succeeded in establishing a free, democratic government. The country is not blessed with deposits of coal, iron, or other minerals. Large-scale manufacturing of goods needed by the population is impossible. They must depend on other countries for imports of automobiles, machinery, and clothing.

Sheep and llamas are raised by Indians who live on the high mountain plateaus of Peru. Along with Christianity, the Spaniards tried to impose the Spanish language on the natives. The Indian masses have remained alienated from the Spanish culture. Luscious sweaters and blankets are created by the women from alpaca and wool on their hand looms, just as their ancestors did. The garments eventually make their way to the fashionable boutiques of Europe and the U.S. where they command high prices.

Where the Andes Mountains thrust downward to the Pacific Ocean, there is no moisture in the air, and even wooden implements and cloth, made hundreds of years ago, are perfectly preserved in the sand. Many Peruvians are small farmers whose families manage to subsist on what they grow. Others are employed on large plantations where coffee, cotton, and sugar cane is produced for export.

Thus Peru and Argentina may develop very differently in the future.

Macrostructure Levels Scoring

Author's Macrostructure of Peru and Argentina Text

Level 1: Topic and Conclusion

        1. label (It's about Peru and Argentina)

        2. a comparison

        3. many factors influence their development

Level 2: Inferred Subtopics

        1. geography

        2. economy

        3. society

Level 3: Subheadings

Peru: 1. land

        2. climate/vegetation

Arg: 3. land

        4. capitol city

        5. farming

        6. natural resources

Peru: 7. farming

        8. natural resources

        9. Indian's livelihood

       10. population

       11. government

       12. education

Arg: 13. population

       14. education

       15. government

Level 4: Details

Peru: 1. mountains

         2. valleys are cold

         3. thin air

         4. on the Pacific

         5. very dry

         6. things are preserved

Arg: 7. large plains

         8. called the Pampas

         9. many cattle ranches

        10. rivers/Atlantic Ocean

        11. capitol has buildings/avenues like Paris

        12. fertile soil

        13. wheat fields

        14. a surplus of wheat and beef is produced

        15. no iron, coal, or minerals

        16. little manufacturing

        17. has to import goods

Peru: 18. farmers grow corn

        19. enough to feed their families

        20. work on plantations/produce coffee & sugar cane

        21. for export

        22. mountains have gold, silver, lead, minerals

        23. government mines for profit

        24. sheep & llamas raised

        25. in high valleys

26. handwoven goods produced

27. sold abroad

28. Indians, descendents of the Incas

29. conquered by the Spanish

30. Indians became poor & oppressed

31. white plantation owners rule

32. Spaniards brought Spanish & Christianity

33. Indians don't speak Spanish/alienated

34. Indians can't read

Arg: 35. few Indians

36. bustling cities/busy people

37. free to publish

38. democracy

Examples of Summaries

6th Grade Summary of Poor Macro/Poor Micro Text

A majority of Peruvians are Indians. They live among the copper, silver, and lead deposited Andes mountains.

Argentina produces more wheat and beef than it can handle.

The Spanish conquerors destroyed the Inca empire, and from then on a small part of plantation owners ruled the country.

In the Argentinian Pampas the soil is very good and needs no fertilizer. All of the Pampas is grassy plain and ranches. The people finally succeeded in establishing a free government. But they must depend on other countries for most of their needs.

In Peru sheep and llamas are raised. Christianity and Spanish language are imposed.

In Peru most people work on small farms or large plantations.

Thus Peru and Argentina may develop very differently in the future.

College Student Summary of Poor Macro/Poor Micro Text

Factors such as geography, government, and people influence the future of a country. Peru is made up of illiterate Indians who are poverty stricken. Argentina, on the other hand, is composed of busy and energetic people who seem to be educated. Peru is blessed with natural resources, especially in the Andes mountains. Argentina has fertile fields but must import a lot of items. Argentina has a democratic form of government, whereas Peru's people are oppressed. It seems as though Peru has the resources but not the people to make use of them and Argentina has the people, but no resources. Therefore, these differences will cause Argentina and Peru to develop differently.

Author   Notes

Requests for reprints should be sent to Eileen Kintsch, Dept of Communication Disorders and Speech Science, Campus Box 409, University of Colorado, Boulder, CO, 80309.

Footnotes

1It should be noted that the college students could hardly be regarded as expert writers. In fact, their summaries contained many exmaples of incorrect grammar and most were unnecessarily redundant and discursive in all conditions. However, in the Poor Macro condition the trend is at least encouraging!

Table 1. Text Propositions and Inferences in the Summaries as a Function of Age Group.

Response Categories

	Text Prop. (no.)	General. (no.)	Elabor. (no.)	Reorder. (no.)	Connect. (percent)
Group					
6th	50.78	5.75	6.16	0.97	6.39
10th	50.63	18.79	10.06	1.88	8.50
CS	47.34	26.91	12.00	2.47	8.15

Table 2. Percent Macropropositions and Details in Summaries and Probe
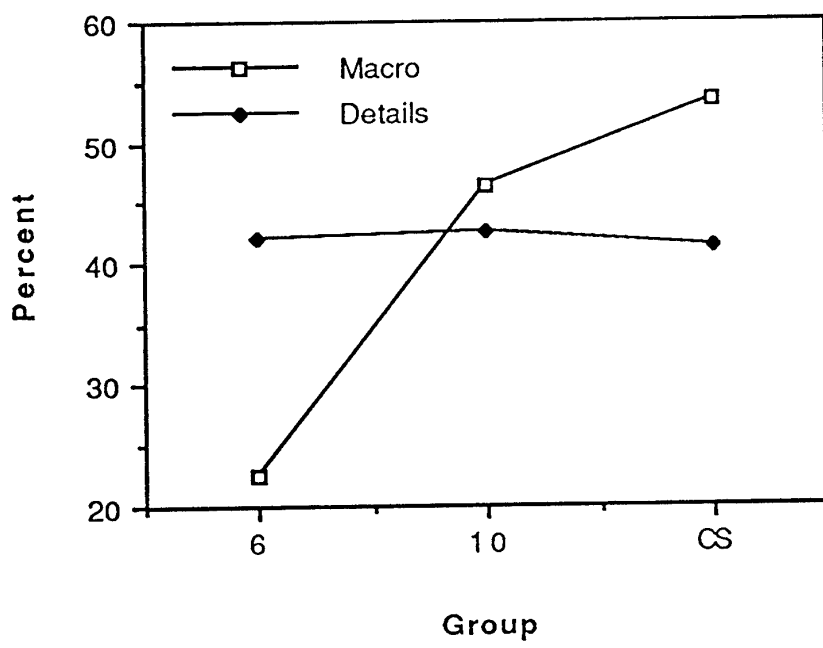Question Responses as a Function of Age Group.

	Summaries			Probe Questions	
	Macro	Detail		Macro	Detail
Group					
6th	22.64	42.10		32.57	9.88
10th	46.33	42.54		54.07	10.03
CS	53.38	41.35		55.89	15.44

Figure   Captions

Figure 1.   Percent text-based statements at each level in the
              macrostructure in summaries.

Figure 2.   Percent macropropositions and details in summaries
              as a function of age group.

Figure 3.   Number of generalizations and reorderings
              as a function of age group and macrostructure.

Figure 4.   Percent macropropositions and details in summaries
              as a function of age group and macrostructure.

Figure 5.   Percent connectives as a function of
              of age group and microstructure.

Figure 6.   Number of elaborations and reorderings as a function
              of age group, macrostructure, and microstructure.

Figure 7.   Percent text-based statements at each level in the macrostructure
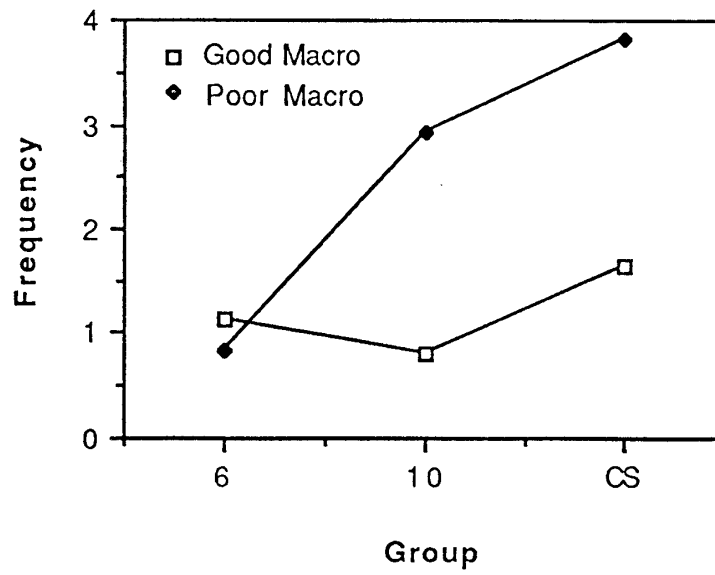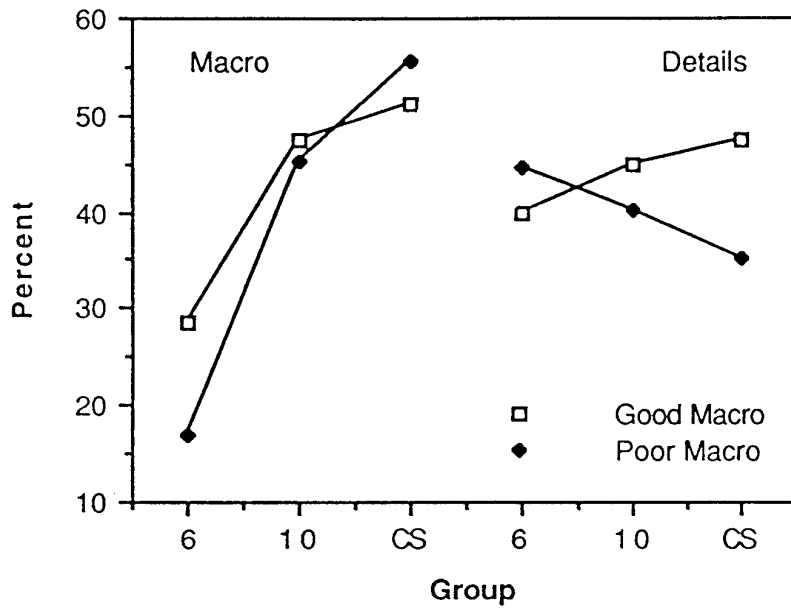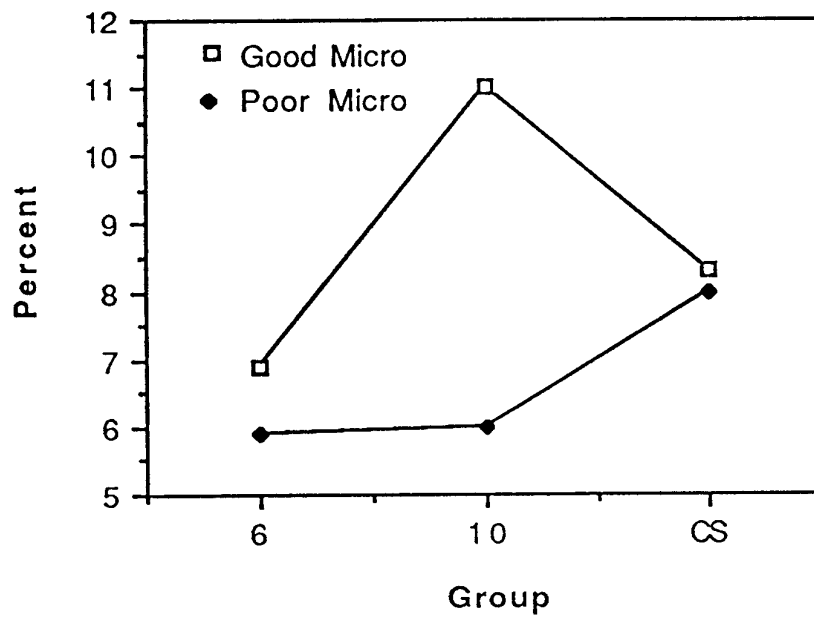   in probe question responses.

Summaries:

Percent vs. Group chart showing Macro and Details series. The Macro series (open squares) rises from about 23 at Group 6, to about 47 at Group 10, to about 54 at Group CS. The Details series (filled diamonds) remains roughly flat near 42–43 across all groups.
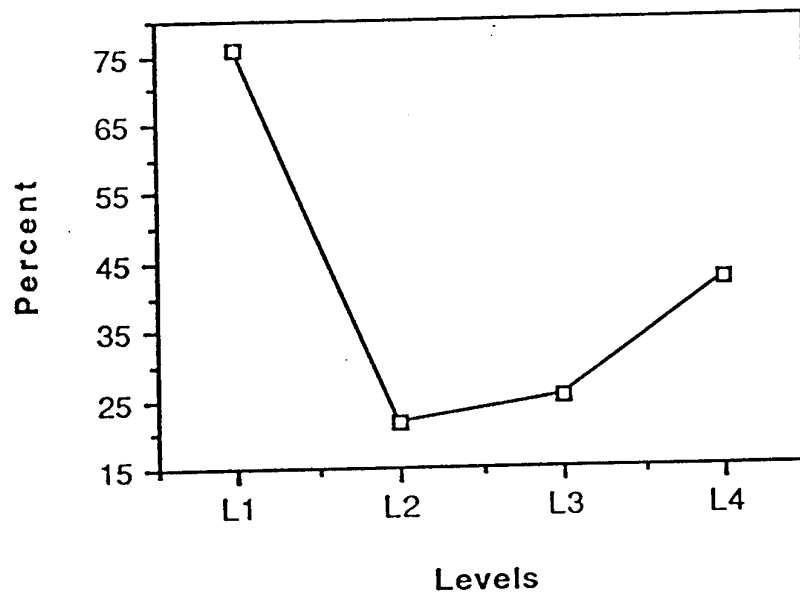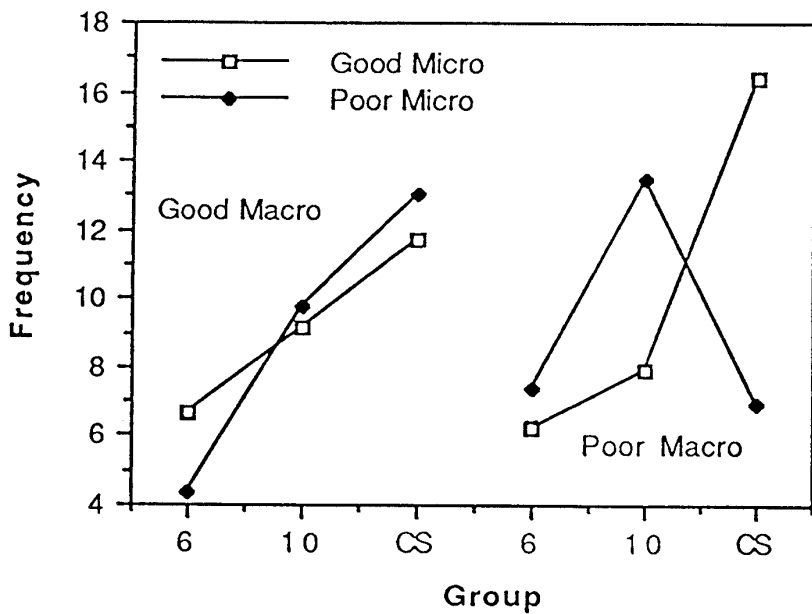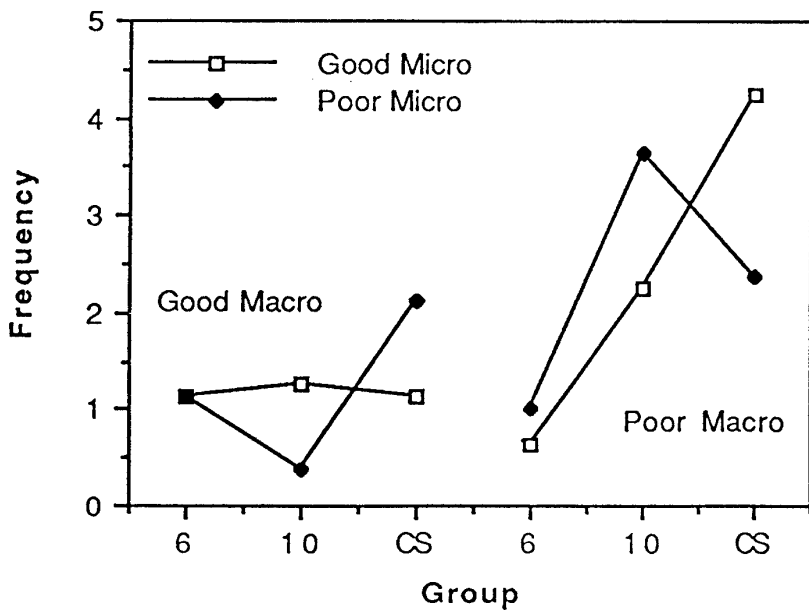
Generalizations:
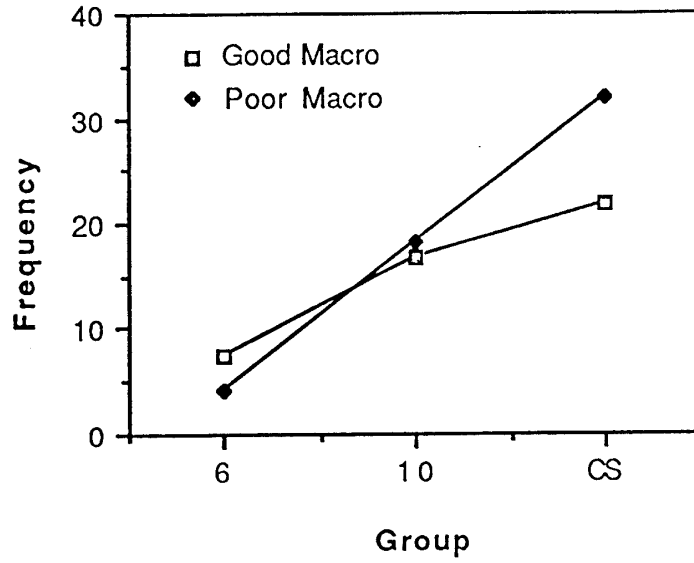
Reorderings:

Elaborations:

Reorderings:

Probe Questions: