ILLINOIS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

PRODUCTION NOTE

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

Technical Report No. 268

THE DEVELOPMENT OF PLANS FOR SUMMARIZING TEXTS

Ann L. Brown
University of Illinois at Urbana-Champaign
Jeanne D. Day
University of Notre Dame
Roberta S. Jones
University of Illinois at Urbana-Champaign

January 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. 268

THE DEVELOPMENT OF PLANS FOR SUMMARIZING TEXTS

Ann L. Brown
University of Illinois at Urbana-Champaign
Jeanne D. Day
University of Notre Dame
Roberta S. Jones
University of Illinois at Urbana-Champaign

January 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 by Grant HD-06864 and a Research Career Development Award to the first author, HD-00111, both from NICHHD.

Abstract

Students from fifth, seventh and eleventh grade, together with college students wrote constrained and unconstrained summaries of stories they had previously learned to criterion. College and older high school students outperformed younger students in their propensity to plan ahead by making rough drafts, in their sensitivity to fine gradations of importance, and in their ability to condense more idea units into the same number of words. The few younger students who planned adequately performed at a level set by college students. Planning, not age per se, was the best predictor of efficiency, although the propensity to plan and age were highly correlated. Under circumstances where summary production is not just the result of "automatic" retention, the ability to work recursively on information to render it as succinctly as possible requires judgment and effort, knowledge and strategies, and is, therefore, late developing.

The Development of Plans for Summarizing Texts

The ability to summarize information is important for understanding and remembering texts and, therefore, the development of this ability in children should be of considerable pedagogical interest. But there are also powerful theoretical reasons why childrens' summarization ability should be examined. Many current theories of text understanding assume, at least implicitly, that a higher-order representation of the super-sentence structure of the text is "automatically" abstracted during comprehension, and it is this macrostructure that guides the production of recall and summarization (Kintsch & van Dijk, 1978; Rumelhart, 1977). In general, comparisons of adults' recall and summarization of texts have been used to support the claim that the same underlying processes are involved in both productions. A summary representation of a story is the natural product of understanding; when asked to recall, subjects have this summary available in memory and can call upon it to serve as the scaffolding of their reconstruction. If this modal theory were correct, then the ability to recall a text would be dependent on the ability to summarize. It follows, therefore, that summarization and recall should be closely linked. There are strong reasons to question this assumption (Johnson, 1978, in press).

Developmental data are particularly pertinent to this argument, for it might lead to the further supposition that children would be quite competent at summarizing texts, a position that would scarcely receive support from the educational literature (Germane, 1921a, b; Stordahl &

Christensen, 1956). We know that under certain circumstances young children recall stories well (Mandler, in press; Stein & Trabasso, in press); and, the relative invariance of story recall patterns over age, ability and cultural milieux (Mandler, Scribner, Cole, & Deforest, 1980; Smiley, Oakley, Worthen, Campione, & Brown, 1977) has tended to support the idea that an "automatic" by-product of comprehension is retention, at least of the main gist of the story. If recall were in fact dependent upon summary, adequate recall would, then, imply an intact ability to summarize the stories. Children who can recall stories should also be able to summarize them. There are two problems with this position: (a) how good is children's recall? and (b) how do we estimate the ability to summarize in a method uncontaminated by amount recalled?

Consider first the question of recall efficiency in children. Most reports of excellent recall have come from studies based on the story grammar approach (Mandler, in press; Mandler & Johnson, 1977; Stein & Glenn, 1978; Stein & Trabasso, in press) where the texts have been ideal stories conforming to the grammar, or from studies using simple picture sequences (Brown & Murphy, 1975; Day, Stein, Trabasso, & Shirey, Note 1). There is ample evidence that given simple, short, well-formed stories that are concerned with familiar content and presented in an hospitable environment, children will generate excellent recall protocols. But what happens if the texts are not so perfect? What happens if the texts are more like those encountered in schools in that they are lengthy, less than well-formed, contain ambiguity, irrelevancy and redundancy, etc.?

There are data indicating that children's processing of less than ideal text materials is not optimal. Children tend to reconstruct the main theme but their recall is somewhat meager compared to that of adults. And, this developmental difference is greatly exacerbated if additional study time is provided. Under these circumstances older children and adults increase their recall of salient gist by the deliberate use of attention-focusing strategies. Younger children tend not to improve upon their original brief recall (Brown & Smiley, 1978). Apparently, then, a bare outline of the main points of a passage may be the relatively automatic result of comprehension, but in order to increase recall beyond that outline, a variety of complex, deliberate activities are needed. These include mechanisms for identifying, selecting and focusing on important elements at the expense of trivia.

Adequate recall of anything more than simple short stories requires both effort and judgment. Recall, if it is to include more than a bare skeleton, demands strategies for concentrating on difficult and important elements; it requires judgment of what to include and what to omit. Similarly, summarizing texts should also entail judgment and effort if more than the barest synopsis is required. Estimations of fine degrees of relative importance must be made and rules for condensation employed. If this were the case, one would predict that the ability to provide an adequate summary of a lengthy text would be a late developing skill.

A problem with estimating children's ability to summarize is that it is essential that we are able to distinguish between the "automatic" by-

product of comprehension, and the deliberate result of judgment and effort. A true summary should be a reduction in length relative to the remembered representation of the text; to summarize implies the ability to condense intelligently what is retained of the gist. If children as well as adults "automatically" extract the main gist when comprehending stories, there is a danger that children will produce a protocol that looks like an adequate summary, not because they apply condensation rules to the remembered gist, but because they produce all the information they have available in memory.

In examining the development of summarization ability it is, therefore, necessary to ensure that one can distinguish between a product that is "all that the child remembers" and one that is the result of judgment and effort. One method of attacking this problem is to ensure that subjects can recall much of the information they are required to summarize. Two ways of accomplishing this are (1) to use simple stories that are well-formed according to a story grammar view and, therefore, lead to excellent recall. This is the procedure used by Johnson (1978, in press) in her developmental study of oral summarization in young children. (2) The second alternative, one adopted here, is to use more lengthy, complex stories and require that the students learn the texts to some criterion before preparing a summary. Under these circumstances it would be possible to examine the students' judgments concerning what elements to include or omit without confounding memory and selection.

Another method of distinguishing deliberate strategic processing from "automatic" memory (Hasher & Zacks, 1979) is to somehow engineer situations

where children will be likely to engage in overt activities that reflect their cognitive processing. Therefore, in this study we examined the development of summary writing when the material to be paraphrased was well known, as well as overt planning activities prior to and during the summarization task.

Method

Subjects. Subjects of four age levels were recruited via advertisements: fifth grade (N = 15, mean CA = 10:11), seventh grade (N = 16, mean CA = 13:4), eleventh grade (N = 15, mean CA = 16:4) and first year college students (N = 11, mean CA = 19:4). In the school samples, approximately half of the students at each grade were female (Grade 5 = 8 male, 7 female; Grade 7 = 6 male, 10 female; Grade 11 = 8 male, 7 female). The college sample was predominantly female (2 male, 9 female). To the best of our knowledge all subjects were naive experimentally and were performing at grade level on basic academic subjects, i.e., students and their parents reported that they were not now receiving (nor ever had received) any special help with reading. The students were attending school at the grade level appropriate for their age, and the students' grade point was A or B.

Materials. Six folk stories[1] were selected from a collection used in several previous prose processing studies. Two of the stories, "The Dragon's Tears" and "How to Fool a Cat," were Japanese children's stories used by Brown and Smiley (Brown & Smiley, 1977, 1978; Brown, Smiley, & Lawton, 1978). The remaining four folk tales (A Test of Skill; The

Squire's Bride; The Father, His Son and Their Donkey; and The Kettle that Could Not Walk) were selected from a collection of such stories because extensive pilot study has shown that they shared all essential features of the Cat and Dragon stories. They were of comparable length (approximately 500 words) and contained the same number of idea units (60). They were of comparable readability (fifth grade, 5.2 - 5.5) as measured by the Dale-Chall index. In addition, they met many of the informal criteria suggested by Kintsch and Vipond (1979) as indicating semantic readability levels: e.g., they took the same amount of time to type and read. Studies of eye movements while reading these stories produce a similar pattern of recessive movements. In prior use of these texts, effects due to story factors have not been reported.

The stories were divided into idea units following a procedure used by Johnson (1970) and Brown and Smiley (1977). Independent groups of between 19 and 27 college students were asked to read two of the stories (randomly selected) thoroughly and then to divide the text into individual units; an individual unit was defined as one that represented a self-contained idea. Ideas could be expressed in as little as one word or as much as several sentences. After division into idea units, each story was retyped with one unit per line, and another group of college students ($\underline{N}$ = 23-41) was asked to rate the importance of each unit to the theme of the story using a four-point scale (for fuller details of the rating procedures, see Brown & Smiley, 1977).

Procedure

Each subject was given two of the stories, randomly selected, and told to take them home and learn them perfectly. They were instructed that perfectly meant that all the ideas of the story, even the little (trivial) ones were to be remembered, but they could remember them in their own words. Practice on a simple story was given to illustrate the difference between gist and verbatim recall. The students were told to take as much time as they needed to learn perfectly and to test themselves periodically to see if they were doing well. We also asked the subjects to keep track of how much time it took them to learn the stories. For the younger subjects, the instructions were given both to the child and the parent.

Approximately one week later the subjects were brought back into the laboratory and required to write down all they could remember of the stories including all the details. After recalling the stories, the subjects were given a break. Then one story was selected randomly to be the text summarized by the subject. The subjects were told that they were to pretend that they were newspaper reporters and had to write a summary of the story for their paper, and that a summary was a short version of the story using the smallest number of words. They were given both a sheet of paper on which to write the summary and scratch paper. They were told that they could write on the story or the scratch paper if they wanted to but they had to write their summary in the space provided. They were told that they could take as much time and do whatever they liked in order to produce the best summary they could. While writing the summary, a copy of the story was also available to them.

After they had finished the first summary (and a break), the experimenter told the children that the editor (person in charge of space) had cut their space because there was a very important story that must be covered. Their task now was to write the story again but in only 40 words. Forty words was selected because the mean number of words used by experts when summarizing these stories was 42. The students were given large sheets of paper with forty spaces at the bottom where they were to write their summary. No explicit instructions concerning the use of the top half of the page (blank) were given, although the students were reminded that they could write on the story, use the scratch paper, and do anything they found helpful in preparing their summary. This was designed to serve as a mild hint that the subjects should make a rough draft before attempting to fill in the spaces. In the final phase, the above procedure was repeated except that the subject's space was cut back to 20 words. The answer sheet contained 20 blank spaces on the bottom and the top half of the page was empty.

After summarizing the story, the subjects were asked to divide the idea units of the second story into four piles corresponding to how important they were to the theme[2]. In all, the children were in the lab for several hours; the time was made less tedious for them by interspersing computer games, etc., between sessions of work. All subjects were paid $5.00 for their participation.

Results and Discussion

All recall and summarization protocols were scored blind by two independent raters who scored for gist recall or inclusion in a summary of each idea unit. Interrater reliability was .94. All recall and summarization data were entered into analyses with stories as a fixed or as a random effect; as this manipulation did not result in any differences in the pattern of results, we report only the analyses with story as a fixed effect. Consideration of the groups' means in the school samples revealed no differences due to sex and, therefore, this factor was not considered further.

Recall Data

Initially we set entry criterion for the summarization phase of the study at 80% correct recall at each level of importance (1-4). This 80% level was set because in a series of studies where high school students engaged in multi-trial (3-6) free recall attempts with these stories, recall accuracy reached asymptote at the 80% level (Brown & Campione, 1978; Brown, Smiley, & Lawton, 1977). However, getting younger subjects to reach this criterion was not as easy as we had expected. Only 8/15 fifth, 0/16 seventh, 7/16 eleventh and 8/11 college students reached the criterion. On inspecting the recall patterns produced, it appeared that there were two main types of subjects. The majority of subjects (65%) managed to reach a criterion of 70% correct (13/15 fifth graders, 6/16 seventh graders, 8/15 eleventh graders, and 10/11 college students). Of the remaining 20 subjects, all but four displayed a pattern that would be characteristic of

a subject attempting to recall the story with only one reading. That is, total recall was below 50% and the pattern of recall heavily favored the most important units (Brown & Smiley, 1977). Accepting 70% correct as an entry criterion, therefore, seemed justifiable given this split. Subjects meeting this criterion differed dramatically from those excluded in that they did show a recall pattern indicative of extra study, and recall was acceptably high at all levels of importance. All analyses reported here are on the subjects reaching the 70% criterion. Parallel analyses of only those subjects reaching the original criterion of 80% produced essentially similar results.

Even with this more lenient criterion, the seventh graders were the most difficult group, with only 38% reaching an adequate level of recall. We have no explanation for this finding except an anecdotal one. The fifth graders were concerned about their performance; their parents reported that they had worked very hard and talked about the experiment a great deal in the learning week; the children themselves seemed honored to be paid as consultants. The initial impression was that the children did not relax until they had finished writing their story recall. Our suspicion is that seventh graders did not work as hard as they needed to. Some support for this anecdote comes from the parents' reports of how long the child spent studying. Parents of fifth graders reported that their children studied for approximately one hour and 40 minutes while parents of seventh graders reported a total of 49 minutes.

A mixed analysis of variance, with Age and Importance Level as factors, was conducted on the recall data. Even though the variance was reduced by entering only those subjects who reached the 70% correct criterion, the Importance Level factor was reliable, $F(3,99) = 24.39$, $p < .001$. The mean proportion correct for level 1 (unimportant) through 4 (most important) units was: .76, .74, .84, .89 respectively. Apparently, it is not easy to remember less important material even if one tries. Of main importance, our selection procedure ensured no age effects in the recall of units from each category.

Summarization Data: Importance Ratings

A particular subject's data were entered into the analyses only if she met two criteria: (a) the previously mentioned 70% correct at all of the four levels of importance and (b) he obeyed the word limitations imposed at each stage, i.e., if called upon to give a 40 word summary he did so. No subject who passed criterion a failed criterion b. Thus, all subjects considered were obeying the basic rules of summary and had reasonable retention of the stories.

In all the analyses of the summarization data the patterns shown by seventh and eleventh grade and college students were essentially similar, with college students showing their superiority by including a few extra units in their summary even though no extra words were permitted. The fifth graders were the aberrant group, but they were aberrant in a manner predictable from prior studies. We will begin by reporting the data for each phase and return to the explanation for the fifth grade pattern later.

The data from the summarization phases of the study are illustrated in Figure 1. Visual inspection suggests that all age groups were obeying the

---

Insert Figure 1 About Here

---

rules of the game and reducing the size of their summary on command. Separate 4 (Age) x 4 (Importance Level) mixed analyses of variance were conducted on each of the summaries: free, 40 word and 20 word. In the free summary condition the main effects of Age, $F(3,33) = 2.98$, $p < .05$ and Importance Level, $F(3,99) = 81.11$, $p < .001$, were reliable as was the Importance Level x Age interaction, $F(9,99) = 3.90$, $p < .001$. Post-hoc tests (Scheffé) revealed that the fifth graders were the aberrant group and we will return to this point later. Seventh and eleventh graders and college students did not differ, all groups showing a clear effect of Importance Level in the selected items. Important units were included in the summaries while trivial units seldom were.

The analysis of the 40 word summary resulted in the same pattern: main effects of Age, $F(3,33) = 4.62$, $p < .01$, and Importance Level, $F(3,99) = 128.34$, $p < .001$, as well as the Age x Importance Level interaction, $F(9,99) = 4.25$, $p < .001$. Again the aberrant group was the fifth graders; seventh and eleventh graders and college students did not differ. For these groups, almost all of the units included in the summaries were level 3 and 4; the unimportant units have dropped out.

The 20 word summary analysis resulted in a main effect of Age, $F(3,33) = 2.86$, $p < .05$, and Importance Level, $F(3,99) = 86.21$, $p < .001$, but no reliable interaction. Post-hoc tests revealed a significant age difference with college students outperforming fifth, seventh and eleventh graders, who did not differ. On this abbreviated version, the pattern of results is the same across ages. Under the constraint of the 20 word summary condition all subjects include only level 3 and 4 units in their summaries.

An overview of the summarization performance of seventh and eleventh graders and college students is that they include in their summaries primarily level 3 and 4 units. While some lower level units (1-2) are included in a free paraphrase, these drop out as soon as space pressure is exerted. Even though the patterns of results are very similar for seventh and eleventh graders and college students, college students' greater efficiency is reflected in their ability to include more idea units in the same number of words.

A word of caution is in order concerning the seventh grade data. As only half the seventh grade sample provided usable data, it could be that the seventh graders who were included were also the brightest students. The fact that they performed approximately at the level set by eleventh graders could be an artifact of this selection bias. This would then be an alternative to the admittedly weak motivational (time-on-task) hypotheses offered here. Note, however, that: (a) Even omitting the seventh grade data we are still left with a comparison of fifth vs. eleventh grade and college, an improvement over the only existing study of summarization in

children, where first and third graders are compared with college students (Johnson, 1978); (b) The seventh grade recall data is equated with that from all other age groups; no age differences in recall cloud interpretation of the summary data. Being able to equate baseline performance is at least as important as some unspecified brightness criteria; (c) As we will see, seventh graders are poor planners and practice the inefficient copy-delete strategy as do fifth graders but unlike the older subjects. Thus, the only difference appears to be the lack of adequate preparation time; and while weak, this is the most parsimonious explanation of the high seventh grade drop-out rate. To be conservative, however, we advise caution in the interpretation of the seventh grade data.

The fifth grade pattern is somewhat different but intriguing in light of prior data from this age group. In a previous study (Brown & Smiley, 1977), fifth graders had been asked to rate the units of these stories for importance. The fifth graders' ratings differed from older children in that they were only able to distinguish level 4 units as more important to the theme than the remaining units.

Differences in rating data have been found to influence the note-taking (Brown & Smiley, 1978), underlining (Brown & Smiley, 1978), and retrieval-cue selection (Brown, Smiley, & Lawton, 1978) of fifth grade subjects. Whereas older subjects show a clear pattern of underlining, note-taking, and cue selection, reflecting the various levels of importance of idea units, fifth graders show a preference only for level 4 units over

all the other levels. This pattern of results has been taken as evidence of a close link between knowledge factors and performance. As the fifth graders only recognize level 4 units as more important, they concentrate extra effort on these units alone when trying to learn these texts. Older students distribute their attention as a function of the finer degrees of importance that they are able to recognize.

The same pattern was found here with the unconstrained summary. When left free to paraphrase in as many words as they wished, fifth graders included many more level 4 units in their summaries than any other level but showed no significant preference for level 3 over levels 1 and 2. However, as space pressure was exerted in the constrained summaries the pattern changed in an interesting way. Constrained to 40 words the fifth graders dropped level 1 units. On the 40 word summary, the fifth graders included significantly more level 4 units than they did level 3 and level 2 which did not differ. The differences between level 2 and level 1 were also reliable. Further limited to 20 words the fifth graders dropped the level 2 units and thus ended up with a pattern indistinguishable from older subjects. This dropping of units as a function of their rated importance is the first evidence, from an extensive series of studies, that fifth graders are in any way sensitive to fine differences in importance between levels 1, 2, and 3 of these particular stories. When severely pressed for space even the younger children are sensitive to fine degrees of importance, a sensitivity that comparable age children do not show when studying the texts (Brown & Smiley, 1977, 1978) or rating them under a

variety of conditions (Brown, Smiley, & Lawton, 1978). This is another illustration of the fact that estimating a child's knowledge or "awareness" (i.e., her metacognitive status) is not a simple task; the degree to which a subject will be judged "aware" depends on the indices used to measure that awareness (Baker & Brown, in press; Brown, in press; Brown, Bransford, Ferrara, & Campione, in press). Only when severely pressed for space do fifth graders indicate that the lower level units of these texts differ in terms of their importance to the theme of the story.

Copy-Delete Strategies

In a series of studies on adolescents' study skills, Brown and her colleagues (Baker & Brown, in press; Brown & Smiley, 1978) have repeatedly found qualitative differences between older and younger high school students. The most common strategy used by fifth-seventh graders, while taking notes and outlining, for example, has been termed the copy-delete strategy. The components of this strategy are (a) read text elements sequentially; (b) decide for each element whether to include or to delete; (c) if inclusion is the verdict, copy the unit more or less verbatim from the text. This general strategy is employed by fifth and seventh grade notetakers (Brown & Smiley, 1978), and it is also applied to the tasks of outlining. Interviews conducted with seventh-eighth grade students concerning their study and research habits again suggest that this is a common method. The students often reported that they copy verbatim from research sources when preparing papers; they had little appreciation of the need to extract the main points and restate them in their own words.

Older students depart radically from the copy-delete ploy in their notes, outlines, book reports, etc.; they rely on paraphrase and condensation rules to combine and rearrange idea units, and to depart from the words actually present in the text. They state the gist in their own words and rearrange the order of items into topic clusters (Brown & Smiley, 1978).

The same pattern was found in the summarization task reported here. The students' summaries were rated on a three point scale of verbatim, near verbatim and paraphrase. Fifth and seventh graders produced 27% verbatim and 57% near verbatim units, while eleventh grade and college students produced 28% verbatim or near verbatim units. The majority of the older students' productions were paraphrased in their own words (69%), while only 16% of the younger children's productions were rated as true paraphrases. The condensation rules used by older students while writing summaries of expository texts are described in more detail elsewhere (Brown & Day, in press).

We also looked at the temporal sequencing of the units included in the summaries. As a crude measure, we determined the proportion of units whose order in the summaries deviated from that in the texts. Fifth and seventh graders rarely departed from text order, with order deviation scores of .08 and .09 on free paraphrase and .08 and .10 on the 40 word summaries respectively. Eleventh graders' and college students' order deviation scores were .24 and .25 for the free paraphrase and .22 and .29 for the 40 word summaries. The age effect for both the free paraphrase, $F(3,33)$ =

10.69, $p < .001$, and for the 40 word summaries, $F(3,33) = 10.68$, $p < .001$ was reliable. Fifth and seventh graders were diagnosed as using the copy-delete strategy in that the majority of the units included in their summaries were verbatim or near verbatim and occupied essentially the same temporal sequence as they had in the text. Eleventh graders and college students departed from the copy-delete strategy by rearranging temporal order and relying much more on rules of paraphrase.

Story Grammars and Other Qualitative Measures

We did consider some other measures of summarization efficiency in addition to the Johnson (1970) rated importance level metric. First we had experts rate the children's summaries; they had great difficulty doing this until we defined efficiency as the amount of pertinent information irrespective of style. This was more successful and the experts' ratings were highly correlated with the simple importance level measure. But unfortunately the experts were not able to add much in the way of refined analyses of what it was they were scoring. Then we developed what we thought were perfect (expert written) summaries, but again scoring against this criterion did not add anything substantial to the picture we obtained using the Johnson measure. In short, the Johnson procedure served its purpose and provided an excellent index of selective attention to gist.

The next consideration was to parse the stories according to the Mandler and Johnson (1977) story grammars and consider the summaries in this light. Subjects of all ages tended to maintain the basic story nodes of the grammar in their recall of the stories. When pressed for space in

the summarization conditions, so that nodes had to be deleted, they omitted subepisodes leaving the basic narrative sequence intact. Pressed even further, they deleted nodes in the order that would be predicted by the grammar -- simple reactions and goals, then endings, particularly endings that were redundant with outcomes (see Johnson, in press; Johnson & Mandler, 1980). These are the types of nodes that are less well recalled (Mandler & Johnson, 1977), and rated least important (Yussen, Mathews, Buss, & Kane, 1980), a nice piece of convergent evidence for story grammars. A similar pattern of results was found for younger children attempting to summarize simple texts orally (Johnson, 1978).

Finally, we considered narrative analyses of the type developed by Omanson (1979), where the content units of the text are classified as central, supportive or distracting based on how the unit is connected to the main narrative line. As might be expected (as the mean rated importance of central units in these stories is 3.30 compared with the mean rated importance of level 4 units of 3.25), in unconstrained summaries subjects above fifth grade maintained central units and some supporting detail while distracting detail was omitted. Further constrained for space, they dropped mainly supporting detail until final summaries contained only the main narrative line. Fifth graders followed the same pattern but were somewhat less efficient than the older students on the free paraphrase. They deleted the distracting details, but their 40 word summaries did contain some supporting details that the older students omitted. In short, the additional analyses confirmed that older students

maintained the main story line, the principal event sequence, the main nodes of the story, as long as possible. Under pressure for space they successively deleted distracting detail, then supporting detail and finally least important nodes, such as ending and internal reactions.

Planning Data

In all phases of the summarization study we provided scratch paper and blank spaces to facilitate any attempts to prepare a rough draft of the summaries before filling in the spaces. Subjects varied in the degree to which they spontaneously prepared rough drafts.

We will report the planning data from the 40 word summary for two reasons. First, this stage did generate the most planning and second, all subjects were recalling predominantly important (level 3 and 4) units; even fifth graders recalled only 3% of level 1 and 14% of level 2 units by the 40 word phase. Thus, recall of units means recall of central units.

We collected all evidence of preparation of rough drafts and then divided subjects into two groups -- those showing clear evidence of intelligent planning and those showing no clear evidence of such activity. Clear evidence would involve having the passage written out and then rewritten, words crossed out, etc. Not clear evidence would be where only a line or two was written, or more commonly, nothing was written and the student proceeded directly to "fill in the blanks." Although this seems a crude distinction, our two independent raters had no problems making the decisions (.98 agreement). A prototype rough draft would consist of a preliminary written version of the passage, often with the number of words indicated, then a rewrite, and sometimes more than one rewrite.

The total number of subjects attempting an intelligent rough draft was 2, 6, 7 and 7 for the four age groups. Considering only those subjects reaching the 70% criterion, significantly more eleventh graders and college students (12/18) planned ahead than did fifth and seventh graders (6/19), $\chi^2 = 4.55$, $p < .05$.

Another interesting index of planning was the informal observation of the experimenter that the younger children not only failed to make a rough draft but also appeared to run out of space, i.e., the children complained that they had no spaces left into which to fit the end of the story. To test this observation we divided the text into two halves so that the halves contained approximately the same number of idea units, distributed evenly over the levels of importance. We then calculated how many units of the first and second half of the story were included in the summaries of subjects who did and did not make rough drafts (referred to as the Plan and No Plan subjects). These data for the subjects reaching the 70% criterion

---

Insert Figure 2 About Here

---

are illustrated in Figure 2. There was a dramatic difference between plan and no-plan subjects at the lower two ages. These younger children who prepared a rough draft showed no effects of position in their summaries. Subjects not planning ahead, however, favored the first half of the story. This confirms the anecdotal report that younger children ran out of space because they failed to leave room for the second half of the story. This

was not true for older subjects or for younger subjects who made a rough draft. A mixed analysis of variance with Age (2: young and old) x Planning (2: plan or no-plan) x Half (2: 1st or 2nd) as the independent variables resulted in main effects for Age, $F(1,33) = 9.97$, $p < .01$; Planning, $F(1,33) = 6.73$, $p < .01$; and Half, $F(1,33) = 3.71$, $p < .06$. Of more importance, the three way interaction of Age x Story Half x Planning was significant, $F(1,33) = 8.61$, $p < .01$. Younger children who failed to make a rough draft tend to run out of space before completing their summary. High school and college students have sufficient control of their activity to enable them to produce an adequate representation of both halves of the story, even without a rough draft.

An important point to note here is that the developmental trend is carried largely by performance on the second half of the story. If one looks only at the proportion of units included in the summary that come from the first half of the story, there is no significant age effect. Fifth and seventh graders perform like college students. However, there is a significant effect of age if one considers the second half data. This point is well illustrated in Figure 2 along with the second point of developmental interest. Although the proportion of planning subjects increases with age, if one considers the summaries of only the planning subjects, there is no effect of age. Fifth and seventh graders perform on a par with college students. The developmental effect is carried primarily by the no-plan subjects. Younger students who fail to plan ahead perform significantly less well on the second half of the text while the

performance of no-plan older students does not show a serial position effect. These data suggest the importance of adequate planning, for without it the younger students perform less well than adults. Of considerable importance, however, is the fact that with adequate planning even the youngest students perform excellently.

Although these data are strongly suggestive, it should be pointed out that they are correlational in nature and there are problems with interpreting such results. The pattern of results just reported suggests that it is the planning strategy that leads to efficiency, and developmental trends showing improvement with age are created by the increased proportion of strategic subjects. This is a reasonable interpretation but as the data are primarily correlational, the interpretation is not that simple. It could be that the young spontaneous strategy users are the brightest children and would perform better than their peers on any task, and on the particular task in question, without the use of strategies. In the present study, the students were all average readers; but even if it were possible to formally partial out ability factors, such as IQ or reading scores, this would not totally bypass the problem. A manipulative study, such as a training experiment where adequate planning is induced in one group and not another, would help confirm the central place of planning in young students' production of adequate written summaries.

General Discussion

Students as young as fifth grade are able to attempt a written summary of lengthy texts but clear developmental trends are still apparent. College and older high school students outperform younger children in their propensity to plan ahead, in their sensitivity to fine gradations of importance in the text, and in their ability to condense more idea units into the same number of words. Under circumstances when a summary is not just a measure of automatic retention, the ability to work recursively on information to render it as succinctly as possible requires judgment and effort, knowledge and strategies. As such, the ability to provide an adequate written summary of a lengthy text is a late developing skill that continues to be refined throughout the school years.

An important finding is the central place of planning. When writing their summaries, fifth and seventh graders who make rough drafts perform at a level set by college students. These data speak to the issue of the relation of metacognition to cognition, albeit somewhat indirectly. When amount or direction of effort is taken as the indication of task knowledge, rather than verbal reports, a very close link between foreknowledge and efficient performance has been found (Brown, in press; Brown et al., in press; Wellman, in press). In this case, the students' tendency to plan could be regarded as a measure of their sensitivity to the task demands, and this sensitivity nicely predicts their summarization performance.

In summary, we would like to emphasize three main points: (a) the gradual emergence of strategic planning; the proportion of fifth and

seventh graders undertaking such activities is quite low; (b) the relationship between effective plans and efficiency; age per se is not the crucial variable, although, of course, age and strategy use are closely related; (c) the close interdependence of strategic action and knowledge; only those students who knew what the important elements were could reflect that importance in their longer summaries. This relationship has also been demonstrated with the tasks of notetaking and underlining (Brown & Smiley, 1978).

These data are compatible with previous findings concerning the development of selective attention strategies for studying texts. During the junior high and high school years, students develop and increasingly fine-tune a battery of serviceable skills for learning from texts. These include: underlining and taking notes on main ideas (Brown & Smiley, 1978), developing macrorules for comprehension, retention, and synopsis writing (Brown & Day, in press; Johnson, 1978), outlining and mapping (Armbruster, 1979), self-questioning (Andrè & Anderson, 1978; Brown, Palincsar, & Armbruster, in press), concentrating on previously missed or difficult segments of text (Brown, Smiley, & Lawton, 1978), and the general propensity of treating studying as a purposive act of attention-directing and self-questioning. All these skills tend to be relatively late in developing because they require a fine degree of sensitivity to the demands of learning from texts. Learners must develop an understanding of (a) available learning activities for directing attention; (b) their own characteristics as learners, including capacity limitations and background

knowledge; (c) the nature of the materials, i.e., text characteristics, including important elements, structural features, etc., and (d) the nature of the criterial task, or the test to which their learning must be put (Brown et al., in press; Brown, Campione, & Day, 1981). The complex coordination of all these factors enables the student to plan, monitor and evaluate her interactions with texts in an economical and efficient manner.

Reference Note

1. Day, J. D., Stein, N. L., Trabasso, T. A., & Shirey, L. A study of inferential comprehension: The use of a story schema to remember picture sequences. Paper presented at SRCD, 1979.

References

André, M. D., & Anderson, T. H. The development of evaluation of a self-questioning study technique. Reading Research Quarterly, 1978-1979, 14, 605-623.

Armbruster, B. B. An investigation of the effectiveness of "mapping" text as a studying strategy for middle school students. Unpublished doctoral dissertation, University of Illinois, 1979.

Baker, L., & Brown, A. L. Metacognitive skills of reading. In P. D. Pearson (Ed.), Handbook of reading research. New York: Longman, in press.

Brown, A. L. The construction of temporal succession by preoperational children. In A. D. Pick (Ed.), Minnesota symposia on child psychology (Vol. 10). Minneapolis: University of Minnesota, 1976.

Brown, A. L. Metacognition, executive control, self-regulation and other even more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), Learning by thinking. West Germany: Kuhlhammer, in press.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. Learning, remembering, and understanding. In J. H. Flavell & E. M. Markman (Eds.), Carmichael's manual of child psychology (Vol. 1). New York: Wiley, in press.

Brown, A. L., & Campione, J. C. The effects of knowledge and experience on the formation of retrieval plans for studying from texts. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), Practical aspects of memory. London: Academic Press, 1979.

Brown, A. L., Campione, J. C., & Day, J. D. Learning to learn: On training students to learn from texts. Educational Researcher, 1981, 10, 14-21.

Brown, A. L., & Day, J. D. Macrorules for summarizing texts: The development of expertise. Journal of Verbal Learning and Verbal Behavior, in press.

Brown, A. L., & Murphy, M. D. Reconstruction of arbitrary versus logical sequences by preschool children. Journal of Experimental Child Psychology, 1975, 20, 307-326.

Brown, A. L., Palincsar, A. S., & Armbruster, B. B. Inducing comprehension-fostering activities in interactive learning situations. In H. Mandl, N. Stein, & T. Trabasso (Eds.), Learning from texts. Hillsdale, N.J.: Erlbaum, in press.

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.

Brown, A. L., & Smiley, S. S. The development of strategies for studying texts. Child Development, 1978, 49, 1076-1088.

Brown, A. L., Smiley, S. S., & Lawton, S. C. The effects of experience on the selection of suitable retrieval cues for studying texts. Child Development, 1978, 49, 829-835.

van Dijk, T. A., & Kintsch, W. Cognitive psychology and discourse. In W. U. Dressler (Ed.), Trends in text linguistics. New York: DeGruyter, 1978.

Germane, C. E. Outlining and summarizing compared with rereading as methods of studying. In G. Whipple (Ed.), The Twentieth Yearbook of the National Society for the Study of Education. Part II, 1921 (a).

Germane, C. E. Value of the written paragraph summary. Journal of Educational Research, 1921, 3, 116-123 (b).

Hasher, L., & Zacks, R. T. Automatic and effortful processes in memory. Journal of Experimental Psychology: General, 1979, 108, 356-388.

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

Johnson, N. S. What do you do if you can't tell the whole story? The development of summarization skills. In K. E. Nelson (Ed.), Children's language (Vol. 4). New York: Gardner Press, in press.

Johnson, N. S., & Mandler, J. M. A tale of two structures: Underlying and surface forms in stories. Poetics, 1980, 9, 51-86.

Johnson, R. E. Recall of prose as a function of the structural importance of the linguistic units. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 12-20.

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

Kintsch, W., & Vipond, D. Reading comprehension and readability in educational practice and psychological theory. In Lars-Goran Nilsson (Ed.), Proceedings of the conference on memory. Hillsdale, N.J.: Erlbaum, 1979.

Mandler, J. M. Representation. In J. H. Flavell & E. M. Markman (Eds.), Carmichael's manual of child psychology (Vol. 1). New York: Wiley, in press.

Mandler, J. M., & Johnson, N. S. Remembrance of things parsed: Story structure and recall. Cognitive Psychology, 1977, 9, 111-151.

Mandler, J. M., Scribner, S., Cole, M., & Deforest, M. Cross-cultural invariance in story recall. Child Development, 1980, 51, 19-26.

Omanson, R. The narrative analysis. Unpublished doctoral thesis. University of Minnesota, 1979.

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

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.

Stein, N. L., & Glenn, C. G. An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), Advances in discourse processes (Vol. 2). Hillsdale, N.J.: Erlbaum, 1978.

Stein, N. L., & Trabasso, T. What's in a story? Critical issues in comprehension and instruction. In R. Glaser (Ed.), Advances in psychology of instruction (Vol. 2). Hillsdale, N.J.: Erlbaum, in press.

Stordahl, K. E., & Christensen, C. M. The effect of study techniques on comprehension and retention. Journal of Educational Research, 1956, 49, 561-570.

Wellman, H. Metamemory revisited. In M. Chi (Ed.), What is memory development the development of? A look after a decade. Human Development, in press.

Yussen, S. R., Mathews, S. R., Buss, R. R., & Kane, P. T. Developmental change in judging importance and critical elements of stories. Developmental Psychology, 1980, 16, 213-219.

Footnotes

This research was supported by Grant HD-06864 and a Research Career Development Award to the first author, HD-00111, both from NICHHD. The authors would like to thank Dannine Cihlar, Sallie Lawton, and Betty Ann Levy for their help in running subjects and scoring the data. Special thanks are extended to Nancy Johnson for her insightful comments on an earlier version of this manuscript and for her generous help in parsing the stories, and in so doing, teaching us to use story grammar analyses. We would also like to thank Rick Omanson for scoring the stories according to his central, incidental narrative line analyses. Author Brown's address is Center for the Study of Reading, 51 Gerty Drive, University of Illinois, Champaign, Illinois 61820.

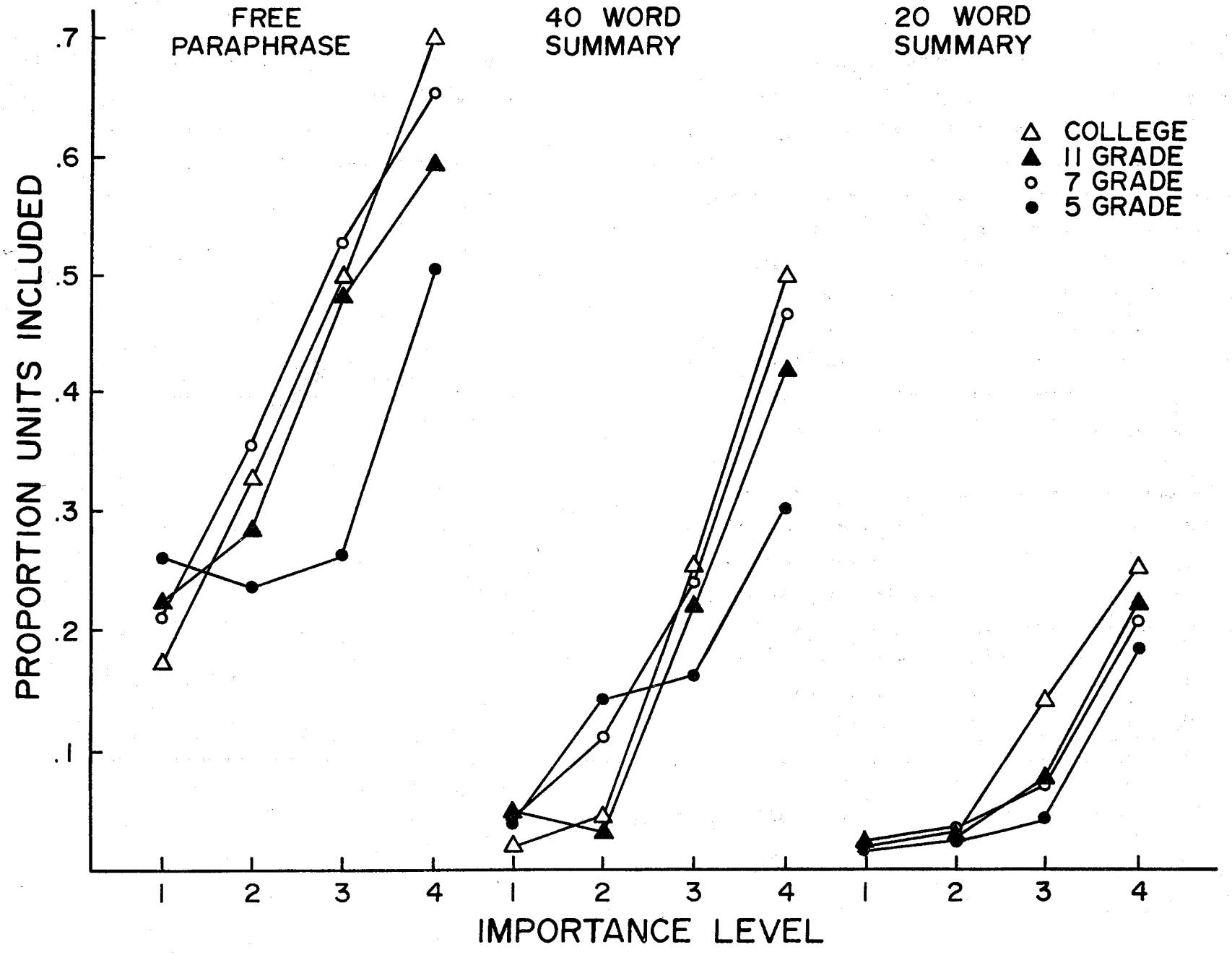
[1]Copies of the stories used in these studies are available from the authors.

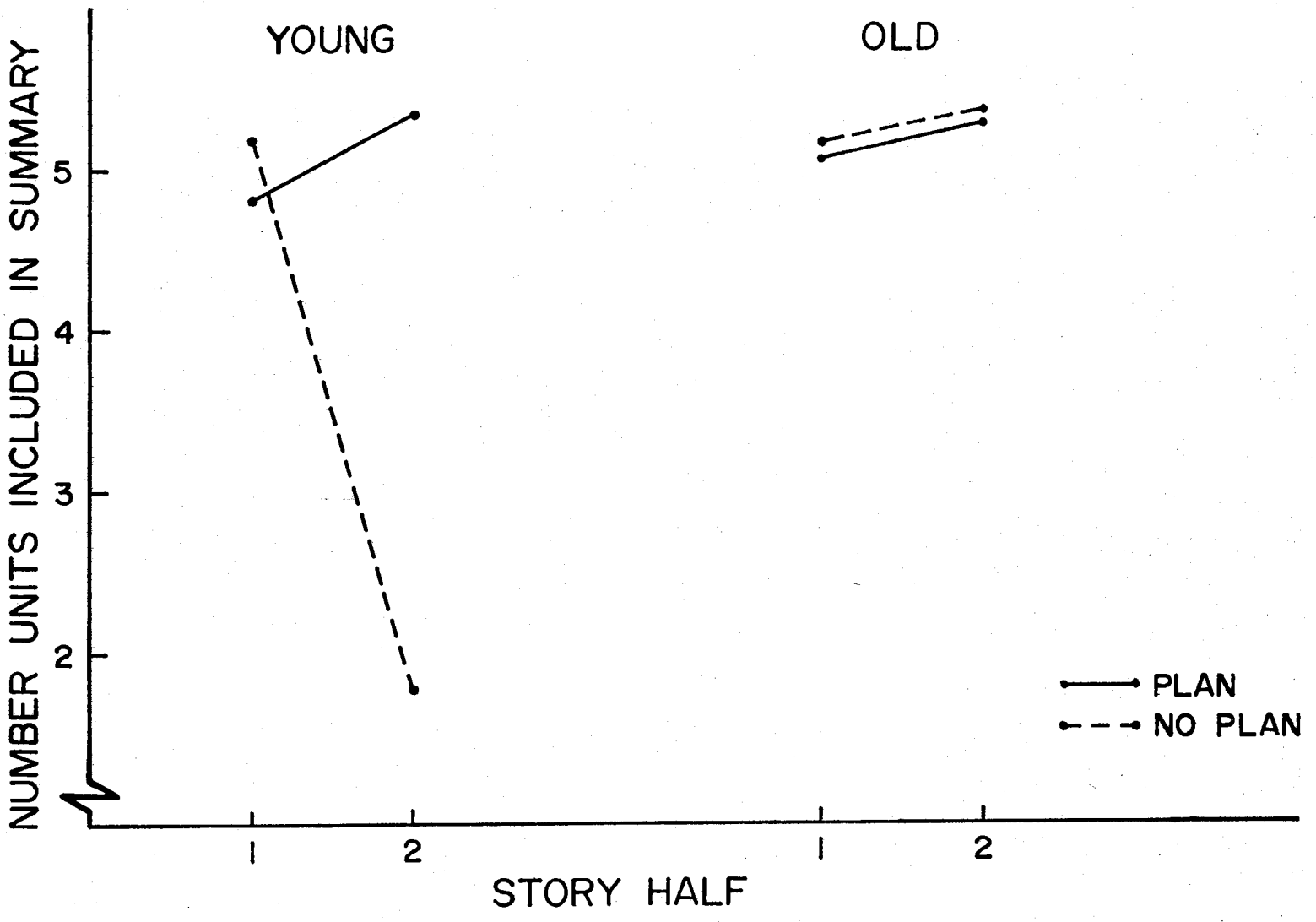
[2]This second task will not be reported. The data served as another replication of the importance rating patterns from students of fifth grade and above on these particular stories. Although we anticipated that the making-piles task would be easier than the Johnson procedure used previously, and would, therefore, change the pattern of results, this was not the case. The pattern of results generated was identical with both procedures (see also Brown, Smiley, & Lawton, 1978).

Figure Captions

Figure 1. The proportion of units included in the summaries as a function of Age and Importance Level.

Figure 2. The number of units included in the summaries as a function of Age, position in the text and planning.


FREE PARAPHRASE
40 WORD SUMMARY
20 WORD SUMMARY
COLLEGE
11 GRADE
7 GRADE
5 GRADE
PROPORTION UNITS INCLUDED
.7
.6
.5
.4
.3
.2
.1
1
2
3
4
IMPORTANCE LEVEL


YOUNG
OLD
NUMBER UNITS INCLUDED IN SUMMARY
5
4
3
2
1
2
1
2
STORY HALF
PLAN
NO PLAN

This page is intentionally blank.