Experimental investigations of children's prose comprehension with special emphasis on the development of skills necessary for reading comprehension are examined in this paper. The research is discussed under four major headings which correspond to specific comprehension skills: identifying main ideas, understanding logical structure, making inferences, and using higher order knowledge structures (the story schema). A preliminary section deals with some of the measurement problems inherent in comprehension research. The paper concludes with a consideration of the research implications for reading comprehension instruction. (Author/FL)
Technical Report No. 102

THE DEVELOPMENT OF PROSE COMPREHENSION SKILLS

Linda Baker and Nancy Stein

University of Illinois at Urbana-Champaign

September 1978


The preparation of this chapter was supported by the National Institute of Education under Contract No. US-NIE-C-400-76-0116. We would like to thank Susan Goldman, Glenn Kleiman, Carol Santa, John Santa, and Tom Trabasso for their helpful comments on earlier versions of this chapter.
The Development of Prose Comprehension Skills

In view of the tremendous role that reading plays in most of our lives, we know amazingly little about the processes involved in reading comprehension. Though this paucity of knowledge may seem surprising, it becomes more understandable when we consider how complex comprehension really is. Comprehension of prose, whether written or oral, involves considerably more than understanding the meanings of individual words and sentences; it also requires that one understand how the ideas expressed in one sentence are related to ideas expressed in other sentences. An even more critical component of comprehension is the knowledge that one brings to the reading situation; in fact, the same passage may be understood in different ways depending on a reader's background. Thus, comprehension involves a complex interplay between the reader and the material.

It is only within the past decade that comprehension has gained widespread attention as a domain of study. Most of the research has focused on adult subjects in order to investigate the comprehension process in mature readers. (See Goetz, in press, and Reder, in press, for reviews of this literature.) Several of the adult findings have stimulated experiments on the development of comprehension, and these developmental experiments will be reviewed in this chapter. The main thrust of the research thus far has been to document the comprehension skills possessed by children of various ages and the changes these skills undergo with age. This information is clearly important to a developmental psychologist, but its value to a reading
educator is more nebulous. The research demonstrates when one might expect certain skills to develop, but it does not provide much insight into how they develop. Though the latter is certainly a critical question to all those interested in comprehension, the research efforts are not yet sufficiently mature to provide the answers. Thus, many of the studies may seem of tangential relevance to reading comprehension instruction. Nevertheless, we feel that because the research has bearing on the development of comprehension, it is, in fact, of practical value. We will discuss these practical implications in the concluding section of the chapter.

In searching for a framework within which to organize our review, we observed that most studies addressed specific skills that contribute to comprehension instead of treating comprehension as a global process. The four most commonly explored skills were identifying main ideas, understanding logical structure, making inferences, and using higher order knowledge structures (most studies have focused on knowledge about the structure of stories). A concern with these four skills is also apparent in comprehension instruction. Typical workbook exercises for beginning readers include finding main ideas, unscrambling sentences in a passage so that it makes sense, making inferences about story characters, and making up or completing stories. Research relevant to each of these skills will be discussed in turn, but the classification is primarily an expository convenience because the skills are highly interdependent. Moreover, we are not suggesting that these are the only skills involved in comprehension nor that they should be studied or taught in isolation.
Some Preliminary Comments on Measuring Comprehension

One of the more difficult problems associated with comprehension research is the choice of an appropriate method for assessment. Just as no one is quite satisfied with existing reading comprehension tests in the schools, so no one is quite satisfied with current experimental methodology for testing comprehension. This problem has been discussed at length elsewhere (e.g., Carroll, 1972; Farr, 1969; Greeno, 1977), but a few comments are in order to acquaint the reader with the techniques that have been used in the research we will be discussing. There are countless variations on the definition of comprehension, and needless to say, the way it is defined influences the way it is measured. Nevertheless, the different definitions share enough common features that comparisons among experiments are meaningful.

Virtually all of the measures of comprehension that experimenters have adopted impose some sort of memory demand on the subjects. Rather than testing for comprehension in the presence of the reading material, as do many standardized tests, experimenters typically present the material, remove it, and then test for memory. The most common memory tests are free recall, probed recall, and recognition. Free recall tests require the subject to produce everything that can be remembered about the material. Such tests are similar to essay questions students receive on exams. Probed recall tests require the subject to provide specific information about the material, and are often in the form of 'wh' questions. These tests are analogous to short-answer exam questions. Recognition tests require the
subject to discriminate statements that are identical to or consistent with the studied passage from related alternatives. Such "objective" tests correspond to the multiple choice or true/false items often found on standardized tests.

The rationale for using memory tests as an index of comprehension is that poorly understood material will not be well-remembered. (Unless it is rote memorized; i.e., many children can recite the Pledge of Allegiance but do not comprehend it.) This rationale received empirical support in studies by Bransford and Johnson (1972) and Thorndyke (1977) where a close relationship was found between the amount of information recalled from a passage and ratings of its comprehensibility. However, one should be cautious in concluding that something has not been understood because it was not remembered. A reader may have good comprehension during reading, but may not be able to remember the material later. Moreover, memory tests of comprehension are plagued with the possibility of a production or response bias. That is, the index of comprehension is based only on the subjects' overt responses; it is possible that something will be comprehended at the time of reading, and remembered at the time of testing, but excluded in the subject's response. Suppose, for example, that you read a story about three little boys named Steve, Mike, and Alan. You are told that Steve is the oldest, Alan is the youngest, and Mike is wearing a blue shirt. On a recall test you might well state that Steve is the oldest and Alan the youngest. But you might leave out the inference that Mike's age is somewhere between that of Steve's and Alan's. You might have left this out because you thought
it obvious, or perhaps because you thought your answer should include only
explicitly stated information. Similarly, you might leave out the fact that
Mike was wearing a blue shirt because, although you remembered it, the fact
seemed trivial and not worth mentioning. As this example demonstrates,
recall tests often provide a faulty index of comprehension, since subjects
are free to decide how much and what information to include. This problem
is better controlled in probed recall and recognition tests, but recogni-
tion tests introduce another bias problem associated with response criteria
(Coombs, Dawes, & Tversky, 1970).

One way to circumvent the potential discrepancy between what the tester
wants and what the reader recalls is to demand verbatim recall (i.e., asking
the subject to recall the exact words of the passage). Most people agree
that this is too stringent a requirement and so subjects are usually
allowed to recall in their own words. However, experimenters often estab-
lish highly subjective criteria in scoring for "gist" recall. In view of
this problem, several researchers have developed models for representing
the semantic content of a passage (e.g., Crothers, 1972; Fredericksen, 1972;
Kintsch, 1974; Meyer, 1975). Since the models represent concepts rather
than individual words, paraphrases and synonym substitutions are permissible
in recall. The models have not been widely adopted by experimenters, how-
ever, because of their complexity. A second type of model that can assist
in scoring decisions is the "story grammar" (Mandler & Johnson, 1977;
Rumelhart, 1975; Stein & Glenn, 1978; Thoodyke, 1977), which will be
discussed in the final section of the chapter. This model, however, is applicable only to a limited class of prose materials: the story.

The major point we wish to communicate is that there is no way to get a complete, unbiased picture of what has been comprehended. However, by using a variety of test procedures, we can hope to obtain a reasonably accurate idea of what the reader has taken away from his or her interaction with a text.

When studying the development of comprehension, additional factors must be taken into consideration. For example, if we want to compare differences in comprehension skills among kindergarten, second, and fourth graders, it makes little sense to ask them to read a passage because the older children can read better than the younger. Thus, in order to avoid confounding differences in decoding ability with differences in comprehension, passages are often presented orally rather than in written form. Another way to avoid this problem is to present narratives in picture format rather than verbal; this approach has the added advantage of sustaining the child's interest level. While there is reason to believe that medium of presentation may affect comprehension (Schallert, Kleiman, & Rubin, 1977), this will not be a focus of our review.

A second problem specific to developmental research is that older children generally remember more information than younger. However, this does not necessarily mean that they comprehended the material better. A number of additional factors contribute to this improved performance on memory tasks, such as improved mnemonic or study strategies, and more
familiarity with testing procedures and task demands. Thus, we should expect to find differences in the amount of information that is recalled by children of various ages; such an outcome is of little theoretical interest or practical value. What is of interest is whether or not manipulations of a particular variable have differential effects as a function of age. For example, it is not very informative to find that older children recalled more from a passage than younger; it is informative to know that the difference was greater when the passage was presented in a disorganized format than in an organized format. Such an outcome would indicate developmental differences in the ability to deal with disruptions in logical sequence.

**Identifying Main Ideas**

Reading comprehension tests abound with questions requiring identification of main ideas. The frequency with which such "main idea" exercises are given to beginning readers is evidence that this skill is regarded as an important component of reading comprehension. Experimental investigations of children's understanding of main ideas have used three general approaches. One approach simply tests for recall of a passage and examines the relative incidence of main ideas in the recall protocols (e.g., Binet & Henri, 1898; Christie & Schumacher, 1975; Korman, 1945, cited in Yendovitskayz, 1971). A second approach is to present children with a passage and ask them to classify the information as to its importance level (Brown & Smiley, 1977). The third approach is to ask children to describe the main idea of a passage in their own words (e.g., Danner, 1976; Mal'tseva, cited in
An early study using the recall approach was carried out by Binet and Henri (1894; excerpted translation in Thieman & Brewer, 1978). Children ranging in age from 9-12 listened to short prose passages of varying lengths and then recalled them. Binet and Henri reported that important ideas were remembered better than less important ideas by children of all ages. In another early study, (Korman, 1945; cited in Yendovitskaya, 1971), children of 4, 5, and 6 years listened to fairy tales and then recalled them. Again, ideas which were related to the theme of the story were more frequently recalled than those which were less related.

A problem common to both studies was that the criteria for deciding the relative importance of ideas were not well specified. The authors presumably used their intuitions to identify the important elements, and it is not clear how much agreement there would be if different opinions were obtained. A more recent study by Christie and Schumacher (1975) attempted to take this problem into account. The authors constructed a 420-word passage that could be divided into 30 "idea units." College students were asked to select the 15 ideas which were most relevant to the theme, and the 15 which were least relevant. The passage was presented on tape to kindergarten, second, and fifth graders who were later asked to recall it. Again, recall was better for ideas judged theme relevant than theme-irrelevant.
Although these results suggest that even kindergarteners are able to differentiate the main ideas from the details of a fairly complex story, this conclusion is suspect. Inspection of the story reveals that the theme-relevant ideas were not simply details; they were deliberately introduced into the story and were noticeably irrelevant (Brown & Smiley, 1977). Thus, even though the kindergarteners differentiated these two classes of information, there is no guarantee that they would be able to do so with "unrigged" stories.

Acknowledging the subjectivity of the previous assessments of importance, Brown and Smiley (1977) adopted a more systematic method for determining structural importance, developed by Johnson (1970). This method first requires that a passage be divided into units that correspond to points at which a speaker would pause. Next, raters are told that the units differ in terms of their importance to the passage and that some of the units can be eliminated without damaging the essence or "semantic cohesiveness" of the text. Units are then classified into four levels of structural importance by first eliminating one quarter of the units judged to be least important to the theme, then the quarter judged next least important, on up to the most important. Although this method lacks a strong theoretical rationale for either the initial parsing of the units or the subsequent ratings, it is a relatively simple way to operationalize importance. Furthermore, it yields a strong predictor of recall; Johnson found that the higher a particular unit was rated in importance, the more likely its recall by college students.
Brown and Smiley (1977) used Johnson's method to determine whether children's recall patterns were also sensitive to the four levels of structural importance. The materials consisted of four non-Western fairy tales of about fifth grade reading level, chosen for their unfamiliarity to most American children. The fairy tales, parsed and rated by college students, were presented on tape to children in third, fifth and seventh grades and were then recalled. The structural importance ratings were a strong predictor of recall: important ideas were more likely to be recalled than less important ideas and all four levels of importance were different from each other in terms of amount recalled. Despite differences in total recall, this same pattern was obtained for children of all three ages, as well as college students. Of most interest was the finding that children as young as eight years were sensitive to fairly subtle gradations in importance. However, six year olds were not able to differentiate the four levels of importance (Smiley, Oakley, Worthen, Campione, and Brown, 1977), although the most important ideas were best recalled, there were no differences in recall of the three lower levels.

The studies discussed thus far suggest that young children recall more of the important information in a passage than the unimportant. However, they provide no indication that young children can actually identify the main ideas of a text. It is possible that differential recall occurs for reasons other than increased attention to important ideas during reading or listening. For example, Brown and Smiley (1977) note that important ideas are usually actions, whereas ideas of lesser importance tend to be
static descriptions. Thus, better recall may result from better memory for events and actions, rather than from identification of the important elements. It has been shown, in fact, that actions are generally better recalled than static descriptions (e.g., Bartlett, 1932; Gomulicki, 1956).

In an effort to determine if children are consciously aware of the differences in relative importance of information contained within a passage, Brown and Smiley (1977) asked students in third, fifth, and seventh grades and college to perform the structural importance rating task. The classifications were then compared to the original ratings done by college students. Third graders were unsuccessful at differentiating levels; their ratings were idiosyncratic, with most units receiving the full range of scores. The fifth graders succeeded in separating the highest level from the other three, which were not differentiated. Seventh graders had somewhat better discrimination: Levels 1 and 2 were differentiated, as were levels 3 and 4, but levels 2 and 3 were not. Only the college students differentiated all four levels.

Although these results suggest that third graders are unable to identify even the most important elements in a passage, it should be noted that this rating task is rather difficult. A number of factors may have contributed to poor performance, among them the complexity of the material; the stories were approximately two years beyond third grade reading level. If the children were unable to comprehend parts of the text, we could hardly expect them to be able to rank the units for structural importance. (This complexity undoubtedly contributed to the low recall scores obtained by the third graders.) Furthermore, the units that the children were asked
to rate were rather small, corresponding, for the most part, to phrases.

It is possible that the children would be more successful at differentiating levels of importance if they dealt with larger meaning units, where the relationship of the part to the whole was more salient.

Bearing in mind that the children's performance would probably be better if the task were simplified, it is interesting to compare the recall results with the rating data. Brown and Smiley found that children from third grade up showed differential recall of the four importance levels, yet not even seventh graders were successful at classifying the units into four levels. One way to account for this discrepancy is to assume that young children's sensitivity to main ideas is below the level of awareness. In other words, selective attention to important elements may be a relatively automatic component of the comprehension process, while overt identification of these elements requires more conscious evaluation of the material. Brown and Smiley suggest this is a problem of 'metacognition'; young children appear to have limited knowledge about their own cognitive processes (Brown, 1975b, Flavell & Wellman, 1977).

It should be obvious that the importance rating task is not the sort of task teachers would use if they wanted to find out if their students could identify main ideas. Although such a task would be useful in revealing whether students could construct a complete outline or efficiently select items for further study, it is too complex to be a good test of comprehension of main ideas. (And, indeed, Brown and Smiley did not intend it to be.)
A few experiments have used more straightforward methods of assessing main idea identification skills, but these studies have a number of weaknesses. In an early study, Mal'tseva (cited in Smirnov et al., 1971) asked children in grades 2, 4, and 6 to compose an outline of a narrative text, highlighting the most important information. The main ideas were extracted by 46% of the children in second grade, 58% in fourth, and 65% in sixth. With increasing age, then, children were better able to discriminate the important from the unimportant. Otto, Barrett, and Koenke (1969) had children identify the main idea in simple, 4-sentence passages. Their instructions were to "make up just one sentence in your own words that says what all the sentences (in the passage) tell you." Only 29% of the second grade children were able to provide an adequate summary statement; most added a considerable amount of detail. In a similar experiment by Danner (1976), children from grades 2, 4, and 6 were asked to identify "the one thing that the sentences in the paragraph tell you about." All children correctly identified two-thirds of the main ideas, and 79% of them identified all. Although older children were more successful, even second graders performed well on this task.

These studies demonstrate that by the time children are in second grade, they have some skill in identifying main ideas. However, the children's abilities may actually be underestimated because of task variables. For example, the children may have been quite successful at extracting main ideas, but they had difficulty producing sentences that adequately expressed them. Moreover, the children may not have understood the rather
cryptic instructions they received in the Otto et al. study. Danner optimized his subjects' performance by giving them a number of orienting tasks, but it is not clear that sufficient practice was provided in the other experiments.

The research is also subject to a criticism raised earlier: There were no explicit criteria for determining the relative importance of ideas. The investigators presumably identified the main ideas themselves and scored the responses for consistency with their subjective standards. Although this is undoubtedly the approach taken by many teachers in evaluating their students' answers, it would nonetheless be desirable to have more objective criteria.

One additional factor to be considered in evaluating the main idea research is that there may be developmental differences in the conception of a main idea. Thus, although the responses did not conform to an adult standard, they may have been consistent with the conception of a main idea at a particular age. This suggestion has received support in a study by Stein and Glenn (1978). Children in first and fifth grade were asked to recall the three most important things that they remembered from a story. The ratings were collected in a successive manner by asking for the first most important thing in the story, the second, then the third. Age differences were obtained in the types of information considered most important. First graders generally focused on the consequences of actions, while fifth graders focused more on the goals of characters in the story.
These results suggest that first graders do have consistent ideas as to what is most important in a story; their ideas just differ from older children's ideas. What remains to be determined is the reason for the developmental shift in importance judgments. One possibility is that the meaning of importance undergoes changes. For example, older children may regard information as important because it helps them organize and remember a text, whereas younger children may consider information important because of its moral value. This is a highly speculative possibility; it is clear that a more thorough investigation of the conception of importance is needed.

In summary, it appears that children as young as five years of age are sensitive to main ideas to the extent that they are more likely to recall main ideas than details. However, it is not clear that this differential recall arises from increased attention to the important elements of the text. This is substantiated by the apparent difficulty young children have in distinguishing important from unimportant information and their less than perfect attempts to summarize main ideas.

Understanding Logical Structure

In addition to extracting main ideas from a passage, an important element of comprehension is understanding how and why the ideas are interconnected. Skill at understanding the logical structure of a text is firmly rooted in prior knowledge of the world. For example, if children do not understand how two events in the physical world are logically related,
Prose Comprehension Skills

It should be noted that some of the studies to be reviewed in this section are not explicitly focused on prose comprehension, but rather the comprehension of logical and temporal relations. They are presented here because they deal with what we believe is an important prerequisite of prose understanding.

Piaget (1926) is responsible for much of the recent interest in sensitivity to logical structure. He reported that in retelling stories, young children frequently mixed up the order of events and expressed causal connections poorly or not at all. He attributed these problems to the child's inability to make use of logical relations. However, the stories Piaget used were lengthy and complex, and so perhaps were difficult to comprehend.

In an effort to clarify Piaget's claims, Brown carried out an extensive program of research investigating children's comprehension and memory for ordered sequences of events. In one set of experiments, Brown and Murphy (1975) presented 4-year-old children with sets of pictures that depicted either a logical sequence of events or an arbitrary sequence. The logical pictures were arranged in either normal or scrambled order. After presentation, the children were asked to reconstruct the ordering of the pictures in each set. Reconstruction was better on ordered logical sequences than on arbitrary or scrambled sequences. This indicates that the children understood the logical structure of the pictures and were able to use their prior knowledge about logical relations to improve memory. Another experiment demonstrated that the same

we can hardly expect them to perceive this relationship in a text.
set of unrelated pictures was better reconstructed when it was accompanied by a narrative which meaningfully interrelated the pictures. Since the pictures themselves were unrelated, the results cannot simply be attributed to correct construction of the order on the basis of prior knowledge alone.

In a similar study, Brown (1975a) tested the hypothesis that the recall failures reported by Piaget were due to excessive memory demands rather than comprehension difficulties. Kindergarten and second graders were shown sets of pictures accompanied by a logically structured or an arbitrarily sequenced narrative. In a third condition, children were instructed to make up their own story to help remember the pictures. After viewing the pictures, the children were asked to recognize, reconstruct or recall the sequences. As before, performance was worse when the narratives were arbitrarily ordered than when they were logically connected. Moreover, the sequences which went with the self-constructed stories were as well-remembered as the logically-structured sequences. Second graders performed equally well on all memory tasks, but for kindergartners, recognition was better than reconstruction which was in turn better than recall. Thus, the more external cues available, the better the performance. This study supports a point made earlier: Recall difficulties do not necessarily reflect failures to comprehend.

Brown's experiments demonstrate quite convincingly that children as young as four years of age understand logical relationships expressed either verbally or pictorially. Moreover, the children are capable of capitalizing on these logical relationships to enhance their memory of the material.
Thus, we have evidence that beginning readers possess the prerequisite skills necessary for comprehending logical structure in prose. It has been shown, in fact, that five year olds are very accurate in recalling the order of events in short stories that are logically organized (Mandler & Johnson, 1977; Stein & Glenn, 1978).

It appears, however, that young children's comprehension is impaired when the order that the events are presented does not conform to a logical sequence. For example, French and Brown (1976) and Homzie and Gravitt (1976) found that preschoolers had poorer comprehension of sentences in which the order of mention was different from the order of occurrence, (i.e., "Before Raggedy Ann calls the doctor, the dog bites the baby"). The disruption was less detrimental when the events were logically rather than arbitrarily related, indicating that the children perceived and benefitted from the logical structure. That young children have difficulty dealing with inverted sequences is not necessarily an indication that they are deficient in a critical comprehension skill. Adults, too, exhibit poorer memory of inverted than forward-order sequences (Baker, in press; Clark & Clark, 1968).

Nevertheless, there appear to be developmental differences in children's ability to deal with disruptions in logical structure. This is reflected primarily in the strategies children use to impose a meaningful organization on the material. For example, Poulson, Kintsch, Kintsch, and Premack (in press) presented four- and six-year-old children with sets of 15 to 18 pictures that depicted a story. A non-memory method of assessing comprehension was used. Children were asked to describe the pictures one by one
as they studied them (after having already viewed the complete set), and their descriptions were compared with adults' descriptions. The pictures were presented in either correct logical order or scrambled, in which case it was extremely difficult for children to perceive the correct sequence.

Most of the descriptions children produced were responses to some feature of the stimulus picture, but they also made responses that could only be derived through an understanding of the story. Twice as many such "story propositions" were produced when the story was intact, which is to be expected since the scrambled pictures did not depict an obvious story. However, many inappropriate story propositions were added in the descriptions of the scrambled stories, indicating that the children were trying to impose a logical structure on the picture set. Six-year-olds did this more frequently than four-year-olds, and they used more inappropriate story propositions when the story was scrambled than they used appropriate propositions when it was intact. Apparently, when the story was well structured, the children felt it would be redundant to add story propositions, but when it was scrambled, additions were needed in order to make the logical structure more apparent. Thus, the six-year-olds seemed to be making up a story as they described the scrambled pictures. In contrast, the younger children frequently reverted to a strategy of labeling the pictures.

Stein (1976) also demonstrated developmental differences in the strategies children use to deal with disruptions in logical structure. In addition, her experiment was a more sensitive test of children's understanding of logical relations among events, in that she used prose stories with subtle disruptions in logical order rather than picture stories with
extreme disruptions. Starting with three logically ordered stories, she systematically distorted them by moving specific statements away from their points of origin. (These statements correspond to "categories" in story-grammar terminology. Further details will be provided later in the chapter.) In one case, the statement was simply inverted with its neighbor, while in the other conditions, it was separated by more statements. Subjects in second and sixth grades listened to the stories and then recalled them.

Although sixth graders recalled more information than second graders, the general patterns of recall were similar. In general, the distorted stories were more poorly recalled than the well-formed stories, with greater effects the further the movement from the original position. Of most interest were the types of reorganizational strategies children used when they encountered a disruption. If the statement was simply inverted, subjects tended to switch it back to its logical position. With larger movements, children often repeated the statement; it was mentioned in the position in which it was heard, but it was also mentioned in the position it should appear. Thus, children remembered the position of the displacement, but they repeated the statement and sometimes added new information to make the story conform to a better structure. The older children were more successful than the younger in reconciling the discrepancies.

In Stein's (1976) experiment, the stories were illogical when they were temporally disorganized; there was nothing to alert the reader that the events had been mentioned out of their proper story sequence. It is possible that if the inversions were marked in the text, there would be fewer
disruptions in recall. This was confirmed in an experiment by Stein and Nezworski (1977); similar types of inversions were used, but markers such as "This happened because ". . ." were included as signals that the order of mention deviated from the order of occurrence. For fifth graders, marked inversions were at least as well recalled as when the information appeared in correct order, and some inversions were actually better recalled. For first graders, some inversions were recalled as well as in well-formed stories, but some were worse. This indicates that young children's comprehension is more dependent upon consistency with a forward-order logical sequence than the older children's. First graders are probably less familiar with temporal inversions as a stylistic device in stories and so are less able to deal with them.

Similar conclusions were drawn by Mandler (1978): She constructed four two-episode stories, and then violated the logical sequence by interleaving statements from the two episodes. Each story began with a common setting, followed by alternating statements from each episode. Subjects in second, fourth and sixth grades, as well as college students, listened to either normal or interleaved stories on tape and recalled them 24 hours later. Not surprisingly, standard stories were better recalled than interleaved. In recalling interleaved stories, subjects frequently repeated the statements in their logical position as well as in their position of mention, a strategy similar to that observed by Stein (1976). Children of all ages were more likely to recall the interleaved stories in their logical sequence than adults; they tended to separate the stories into discrete
episodes, whereas the adults recalled the stories in their order of presentation. Handler attributes this reorganization to a lack of familiarity with discrepant structures; in order to remember a story, children need to make it conform to a logical sequence.

These studies have shown that children are sensitive to logical structure in stories, since deviations lead to decrements in recall. Furthermore, it is clear that children begin to develop strategies for dealing with the deviations by the time they are in first or second grade, as evidenced by their attempts to reconstruct a logical sequence. Thus, these studies are further evidence that skill at understanding how and why ideas are interconnected within a story develops very early, probably before the child has begun to read.

Up to this point, our discussion has focused on children's sensitivity to logical structure in picture and oral narratives. Awareness of logical structure in expository prose is also an important concern, but few studies deal with this type of text. Danner (1976), however, has carried out an initial investigation. He constructed two short passages containing four topics related to an overall theme. In the organized versions, each paragraph dealt with one topic, while in the disorganized versions, each paragraph contained sentences about different topics. Children in grades 2, 4, and 6 listened to the taped passages and subsequently recalled them, with each subject hearing an organized version of one passage and a disorganized version of the second. The amount of text recalled was greater for the
organized than the unorganized versions, and older children recalled more than younger children. In the organized versions, all children tended to group together those ideas that were related to a particular topic sentence; however, developmental differences in grouping strategies were observed with unorganized texts. Older children reorganized the statements to conform to the logical, topical grouping, whereas younger children did not.

After the recall task, the children were tested for their understanding of logical organization: They were asked which passage was more difficult and why; they were asked to state the differences between the organized and disorganized passages, and they were asked to group a random arrangement of sentences into their topical groupings. On all tasks, older children performed better than younger children, suggesting differences in the awareness of the organization that can be built into text material. For example, all children reported that the disorganized passages were more difficult, but only the older children could show the experimenter how the two passages differed or could actually state that one passage was "mixed up" and the other in "the correct order." Furthermore, older children could more easily group sentences in a passage around specific topic sentences.

Danner's results show an interesting parallel with Brown and Smiley's (1977) findings. Whereas all children appeared to be sensitive to discrepancies in logical structure as reflected by amount recalled, only the older children were able to explain why the passages differed in difficulty. Again, we see evidence of a metacognitive deficit. The results also
invite the speculation that children develop an understanding of logical structure in expository prose at a later age than they do in stories. This could result because children are exposed to narratives from the time they first begin to understand language, while experience with expository text is infrequent before third grade.

In conclusion, comprehension of logical structure is an early-developing skill. Children's knowledge about logical relationships and structure greatly influences their memory for prose material. Those passages that are organized according to an underlying logical structure are better remembered than arbitrarily sequenced or disorganized passages. The studies reviewed, however, illustrated that there are developmental differences in the skills brought to these tasks. These differences seem to be related to children's awareness that logical structure has a facilitative effect on memory. Older children were more flexible and competent in using active strategies to increase memory for disorganized material. An important area of future investigation is the process by which this flexibility and awareness of logical structure develops.

Making Inferences

In order to understand the main ideas of a text and perceive their interrelationships, it is often necessary to bring in information that is not explicitly presented in the text. Many of the things readers need to know to comprehend prose are not explicitly stated; therefore, they must be able to draw upon prior knowledge of the world to make inferences.
and fill in "gaps" in the flow of ideas. That comprehension involves an interaction between the incoming information and what a person already knows has been persuasively argued by Bartlett (1932). By providing numerous examples of prose recall protocols, he demonstrated that meaning is not inherent in a text but must be constructed by the reader, and so may differ depending on experience, attitudes and context.

Bartlett's ideas have been influential in stimulating research on the role of inferences in prose comprehension. Although most of the studies have used adult subjects (see Bransford & McCarrell, 1975), developmental psychologists have also become interested in the problem. Much of this work has been reviewed thoroughly elsewhere (Paris, 1975; Paris & Lindauer, 1977; Trabasso & Nicholas, 1977), so our discussion will be relatively brief.

We would like to note at the outset that the research on children's inferencing skills leaves much to be desired. In many studies, it is not clear that the children's performance can even be attributed to the use of inferences. Moreover, the most commonly used experimental task is far removed from normal reading situations. Nevertheless, the work deserves mention, if only to show how many questions are still unanswered.

One of the most frequently tested hypotheses emerging from Bartlett's (1932) work is that people construct an integrated semantic representation as they read or listen to prose and that as a result of this integration, it is sometimes difficult to distinguish the actual text content from inferred information. This hypothesis was tested developmentally
by Paris and Carter (1973), after it received support in an adult study by Bransford, Barclay, and Franks (1972). The materials in both experiments consisted of sets of three related sentences, two premises and one filler. An example is:

- The bird is in the cage. (premise)
- The cage is under the table. (premise)
- The bird is yellow. (filler)

The two premise sentences allow one to infer the transitive relationship, "The bird is under the table." Of critical interest is the extent to which subjects falsely indicate that this true inference had been a member of the acquisition set. In addition to the true inference, recognition items included a true premise ("The bird is in the cage"); a false premise ("The cage is over the table"); and a false inference ("The bird is on top of the table").

In the Paris and Carter study, seven sets of sentences were read aloud to children in second and fifth grades. After a 5-minute delay, the children were given the recognition statements and were asked to decide if they were exactly the same as those studied. Although second graders made more errors than fifth graders, their response patterns were similar. Children in both grades consistently made errors on true inferences; in fact, they were as likely to identify true inferences as "old" as they were to label true premises "old." The children were considerably more accurate in labeling both false premises and false inferences as "new."
These data led Paris and Carter to conclude that children, like adults, construct the semantic relationships among ideas and integrate them in the representation stored in memory; this creates difficulty discriminating inferred from explicit information. Brown (1976b) and Paris and Mahoney (1974) reported similar results using pictorial materials. In all studies, the fact that even the youngest children had difficulty recognizing true inferences as "new" was taken as evidence that the ability to make inferences develops relatively early.

However, a number of factors cast doubt on the conclusion that the children were in fact drawing inferences. For example, Trabasso and Nicholas (1977) suggest that the children may have had a loose decision criterion; that is, they said "old" whenever a statement was semantically consistent, even if it could be discriminated from an actual premise. A second problem arises from the fact that the false statements on the recognition task introduced new relational terms while the true statements retained the original terms (Trabasso & Nicholas, 1977; Thieman & Brown, 1977). Thus, it is possible that children falsely recognized true inferences as "old" items because the relational term was the same and not because they had made the appropriate inference. Some support for this alternative explanation has been provided by Thieman and Brown (1977). Finally, young children are notorious for their bias to respond "old" to items on recognition tests. This bias is particularly a problem when the data of primary interest are incorrect "old" responses.
A recent study by Kail, Chi, Ingram, and Danner (1977) provides somewhat better evidence that children can and do make inferences. The earlier paradigm was modified by having children decide if the test sentences were consistent with, rather than identical to, the stories. This modification eliminated reliance on false recognition errors as an index of inferencing, since responses are correct if true inferences are classified as semantically consistent. Materials were similar to those used by Paris and Carter (1973), except that some of the 3-sentence stories allowed contextual, rather than transitive, inferences. For example, "Mary was playing in a game. She was hit by a bat." invites the inference, "Mary was playing baseball."

Children in second and sixth grades read the sentences aloud from slides, controlling presentation times themselves. After the presentation of each story, subjects received one premise and one inference question. Children at both grade levels showed greater than chance accuracy on all types of questions, and second graders were comparable to fourth. Of most interest was the fact that subjects frequently judged true inferences to be semantically consistent, while correctly judging false statements inconsistent. Thus, this study strengthens the earlier claim that even the younger children made inferences. Furthermore, it shows that they have the ability to make "gap-filling" inferences; i.e., supplying the omitted information that the game was baseball, as well as the "text-connecting" inferences that establish intersentence relationships. It is much harder to argue that the contextual inferences were simply an artifact of the testing procedure.
In all of the studies discussed thus far, the investigators concluded that the inferences were made during initial exposure to the story and stored in memory along with the explicit information. However, it is possible that inferences were not made until the time of test, prompted by the recognition statements. Thus, the data do not indicate that children made inferences during reading but simply that they can make inferences. Of course, knowing that children can make inferences at all is really of most importance. Besides, children should not be encouraged to make all possible inferences as they read, but only those which are necessary. It is not clear that they must make the transitive inference, "The bird is under the table," in order to understand "The bird is in the cage. The cage is under the table."

In view of the limitations of his earlier work, Paris (Paris & Upton, 1976) provided a more sensitive test of children's ability to draw inferences from prose. The materials consisted of passages that were seven or eight sentences in length, as opposed to the simple sentence or picture sets used previously. The passages described behaviors and incidents familiar to young children (e.g., raiding the cookie jar). Eight yes/no probe questions were constructed for each passage, half of which required inferences and half tested memory for verbatim information. The required inferences were of two basic types: those that could be made from single lexical items, (e.g., inferring that scissors were used to cut some paper), and those that depended on contextual relations within and between sentences (e.g., inferring that a child who tried to help a wounded bird liked to take care of animals).
Subjects were children in grades K-5, who listened to each story as it was read aloud and then answered the eight questions. The older children made more correct responses than younger children on both verbatim and inferential questions, but the difference was greater on inferences. Further analysis of the data revealed that the developmental improvement in making contextual inferences was not simply due to better memory of the stories (although the lexical inference improvement was). This result led Paris and Upton to conclude that children's inference-making skills do improve with age, contrary to Paris' earlier conclusion (Paris & Carter, 1973).

A second experiment by Paris and Upton (1976) examined the relationship of performance on the probe task to a subsequent test of free recall. Recall accuracy correlated highly with the ability to draw contextual inferences at each grade level, and this correlation increased with age. The authors concluded that inferencing enhances recall and that the older the child, the more recall is improved. Although this conclusion is intriguing, it should be regarded as tentative: The correlation does not indicate that inferencing caused improved recall, but simply that the two were somehow related.

A few recent studies have provided perhaps the most unambiguous evidence that young children can draw inferences from prose material. These studies have all used a questioning technique specifically designed to elicit inferences. Brown, Smiley, Day, Townsend and Lawton (1977) presented children in second, fourth, and sixth grades with passages that could be interpreted with respect to a previously-provided framework. A series of probe questions
indicated that the children had accessed information from the orienting framework to aid in the comprehension of the story. (See Levin, this volume, for a more thorough discussion of the study.) Stein and Glenn (1978) and Omanson, Warren, and Trabasso (1978) also found that young children could draw inferences from stories. Although the inferences were not always those an older child or adult would have made, it was clear that the children accessed their previous knowledge in dealing with the new material.

In summary, despite weaknesses in many of the experiments we reviewed, the available evidence is sufficiently compelling to conclude that children can and do make inferences about prose material, calling upon their general knowledge of the world to supplement explicit information. Several issues remain to be clarified, however, such as children's awareness of the inference process, the conditions under which inferences are made, and whether or not inferences influence memory.

**Using Higher-Order Knowledge Structures**

Throughout this chapter, we have argued that comprehension involves an interaction between the reader's background knowledge and the text itself. Prior knowledge plays a crucial role in all of the skills we have discussed: extracting main ideas, understanding logical structure, and drawing inferences. When using these skills, specific knowledge may be brought to bear on particular segments of text; for example, we access our knowledge about tools that can be used for cutting paper to infer that "scissors" was the implied instrument in the sentence "She cut the paper." Situations also arise where more generic knowledge can be used to enhance comprehension.
For example, people have generalized knowledge about stories, with specific expectations about their structural components. Similarly, people often have general knowledge about the structure of reading materials specific to their field, i.e., journal articles. Such organized collections of knowledge are known as "higher order" knowledge structures or "schemata." Schemata are thought to facilitate comprehension because they can be used as an organizing framework within which to integrate incoming information. (See Anderson, 1977, & Rumelhart & Ortony, 1977, for further discussion of the role of schemata in comprehension.)

Within the past few years, several investigators have attempted to describe the higher order structures that are used to encode, represent, and retrieve information in stories. Attention has focused on the story because of the regularity in its underlying logical structure. That is, despite variations in content, linguists have observed a stable organizational pattern governing the types of information and logical relations that exist in most stories (e.g., Colby & Cole, 1973; Levi-Strauss, 1955; Prince, 1973; Propp, 1958). Whereas linguists have been concerned primarily with the structure of the stories per se, psychologists have been more interested in the knowledge people have about the structure of stories. This knowledge has been described in a number of different grammars for stories (Handler & Johnson, 1977; Rumelhart, 1975; Stein & Glenn, 1978; Thorndyke, 1977). Despite some differences in terminology and degree of elaboration, the major characteristics of the grammars are similar. The Stein and Glenn (1978) grammar will be summarized here for illustrative purposes.
Table 1 provides an example of a simple story that has been partitioned into categories, the basic units of analysis in the Stein and Glenn (1978) grammar. The story is considered well-formed because it contains all of the requisite categories, arranged in their correct logical sequence. A simple story can first be broken down into two parts: a setting category plus an episode structure. The setting begins the story with the introduction of a protagonist and normally includes information about the social, physical or temporal context of the story. The episode is the primary higher order unit of analysis and consists of five categories of information. These categories serve particular functions in the story and occur in fixed temporal sequence. The initiating event category contains an event or action that changes the story environment. The major function of this change is to evoke the formation of a goal. The goal is included in the internal response category. Internal responses also include affective states and cognitions, and they serve to motivate a character's subsequent overt behavior. Overt actions that are directed towards goal attainment are classified as attempts. The result of an attempt is the consequence, which marks the attainment or non-attainment of a goal. The final category is the reaction, which can include either a character's response to the consequence or broader consequences of the goal attainment.

In reality, few stories have a structure as simple as the one described; most stories contain many episodes, and these may be connected by various types of logical relations. Similarly, stories may also contain incomplete episodes, where one or more of the basic categories is omitted.
In such cases, it is assumed that the reader infers the information contained in the missing category. However, if too many categories are missing, and/or the logical connections are vague, people will not be able to construct an adequate representation of the story. Such stories are not considered "well formed."

This brief description of a story grammar is admittedly over-simplified due to space limitations. The main point we want to convey is that there are rules governing the kinds of information that should appear in a story and the order in which this information appears.

A number of recent experiments have tested predictions about story comprehension and memory based on the grammars (e.g., Glenn, 1978; Mandler, 1978; Mandler & Johnston, 1977; Rumelhart, 1975; Stein, 1976; Stein & Glenn, 1977a, b, 1978; Stein & Nezworksi, in press; Thorndyke, 1977). This work is discussed in detail in Stein (in press) and the interested reader is referred to that source. One of the major conclusions that has emerged is that knowledge of the structure of stories is critical to an understanding of stories. Therefore, it is important to study the development of this knowledge and its role in children's comprehension.

A straightforward way to assess children's knowledge of story structure is to ask them to produce a story. If they do have knowledge about the kinds of information that belong in stories, then this information should appear in their constructions. Stein and Glenn (1977a) provided kindergarten, third, and fifth graders with story settings and asked them to finish the stories. The children's stories were classified according to their
structure, which ranged from simple descriptions through complex episodes. The more sophisticated structures were characterized by their inclusion of purposive behaviors and increasingly well-specified motives and goals. Only about half of the kindergartners’ stories contained purposive behaviors, while two thirds of the third graders’ stories were purposive, as were almost all of the fifth graders’. Thus, there was a clear developmental progression in the logical complexity of the stories, presumably reflecting increasing knowledge of the constituents of a well-formed story.

Leondar (1977) also found a strong relationship between age and the structural complexity of stories produced by children ranging in age from 3 to 16. Similarly, Sutton-Smith and his colleagues (Botvin & Sutton-Smith, 1977; Sutton-Smith, Botvin, & Mahoney, 1976) reported high correlations between age and several hypothesized levels of structural complexity in the stories constructed by children from 3 to 12 years of age. Although all of the investigators used different indices of structural complexity, they observed strikingly similar developmental patterns. Perhaps of most importance is the common observation that children as young as four and five years of age were capable of constructing well-formed, purposive stories. This finding conflicts with Piaget’s (1926) claim that children lack the cognitive structures to produce a coherent story before the age of seven or eight. All of these experiments suggest that children acquire knowledge about story structure at a very early age and use it to guide their story construction.
The extent to which such knowledge influences comprehension and memory of stories has also been investigated. The basic paradigm is to present children with short stories and ask for recall. The primary focus is on qualitative aspects of recall, rather than quantitative; in other words, researchers are more concerned with the kinds of information children remember from stories than the overall amount. Stein and Glenn (1978) presented first and fifth graders with children's stories that had been analyzed according to their grammar. The older children recalled more than the younger, but recall of specific statements was stable over grades. Some categories were more salient to the children than others, as indicated by their frequency of recall. Major settings were best recalled, closely followed by initiating events and consequences. Internal responses were poorly recalled, except when they contained goal statements. The only consistent developmental difference was that fifth graders recalled more internal responses than first graders. This parallels the increasing emphasis on motivations reflected in children's story construction and importance ratings (Stein & Glenn, 1978). [We should point out that this trend is not specific to stories, but appears in children's understanding of many types of social interactions (Flappan, 1968).]

Stein and Glenn (1978) also examined information that had not been contained in the original stories but was introduced in recall. More intrusions were made by fifth graders than first, and the intrusions frequently belonged to the internal responses and attempt categories. The fact that internal responses were poorly recalled would lead one to believe
that children are insensitive to psychological states of the characters; however, the high proportion of internal response intrusions indicates that this is not so. In fact, a second experiment by Stein and Glenn (1978) directly probed children's perceptions of causality in the stories, and showed that all children regarded a character's intentions and motivations as the primary cause of the consequence. Responses to a series of "why" questions revealed that even the first graders had good comprehension of the logical relationships existing among the story categories.

Handler and Johnson (1977) also examined qualitative aspects of story recall using four short stories analyzed according to their grammar (Stein and Glenn's terminology will be used here since it has already been introduced). Children in the first and fourth grades, and college adults, listened to and recalled the stories. Adults recalled more information than fourth graders, who in turn recalled more than first. Age differences were observed in the amount recalled from specific categories, but the patterning of recall was similar. Settings were best recalled by the first graders, closely followed by initiating events and then consequences. Recall was progressively worse for attempts, reactions, and internal responses. Fourth graders had the same ordering of category recall except that attempts were as well recalled as consequences. Adults recalled attempts, settings, initiating events, and consequences equally well, but reactions and internal responses were still significantly worse. These commonalities suggest that young children are sensitive to the same
structural components in stories as adults. The results are consistent with those of Stein and Glenn (1978) in showing differential recall of specific categories. It appears that story grammar analyses can predict what information will be remembered on the basis of its structural role in the story.

A major prediction derived from story grammar analyses is that stories which conform to the prototypical structure will be better remembered than those that do not. Stein and Glenn (1977b) tested this hypothesis by examining the effects of category deletions on children's story recall. They constructed four stories that contained all six categories specified by their grammar and created five variations by deleting one category from the episode. Children in first and fifth grades listened to and recalled either well-formed stories or their structural variants. Fifth graders recalled more than first, but in general, the category deletions did not have the anticipated disruptive effects on recall. However, for both grades, recall was disrupted when the initiating event of the story was deleted, and first graders showed decreased recall when the consequence was deleted.

An analysis of the intrusions in recall proved informative. Fifth graders made more inferences than first graders except when the stories were well formed or when the reaction was deleted. There were more inferences relative to the well-formed story when the initiating event, attempt or consequence was deleted, but no increases with deletions of the internal response or reaction. It is interesting to note that it is when the most frequently recalled categories (initiating events and consequences) are
deleted that most new information is added to recall. Similarly, the deletions of these categories produced the largest decrement in accurate recall. The added information was often of the same category type as that which was deleted; that is, if an initiating event was deleted, children would infer one; if a consequence was deleted, a new one would be inferred. This study provides further evidence that young children do have knowledge of story structure and that they use that knowledge to make deviant stories conform to the norm. Nevertheless, developmental differences were apparent in the skill with which gap-filling inferences could be made.

A related prediction of story grammars is that comprehension and/or memory will be impaired if the presentation order of a story violates the prototypical sequence of categories. Since a disruption in category sequence produces a disruption in the logical flow of ideas, it is intuitively clear that this prediction would be supported by empirical test. In fact, we have already described the studies which have been undertaken as specific tests of this hypothesis in the section on logical structure (Handler, 1978, Stein, 1976, Stein & Nezworski, 1977). To reiterate, these studies demonstrated that young children are sensitive to disruptions in story sequence, as reflected by decrements in recall and attempts to reorganize the story to conform to a more logical structure.

In summary, it appears that knowledge about the structure of stories develops during the preschool years. Most children's exposure to stories begins before they can even talk, so it is not surprising that a story
schema is acquired quite early. The schema goes through refinement during the elementary school years, with an increasing focus on internal goals and responses. Several studies have provided evidence that children, as well as adults, benefit from the organizing framework of the story schema. Story grammars have been constructed to describe the schema and are useful as an approach towards understanding the comprehension process. The grammars offer a model of the strategies people might use when reading or listening to a story that enable them to encode information efficiently.

Conclusions:
What Can Researchers Tell Educators that they Don't Already Know?

As we cautioned at the beginning of the chapter, the relevance of many of these experiments to comprehension instruction is far from obvious; nevertheless, we claimed they were of practical significance. We will now defend this claim, but at the same time point out the limitations of the research and directions for further study.

It is undoubtedly true that many of the experiments we reviewed simply confirmed what reading teachers have always known: under the right conditions, young elementary school children can identify main ideas, understand logical structure, make inferences, and use knowledge about the structure of stories. Perhaps teachers would feel gratified to know that their intuitions and classroom observations have been supported experimentally, but they would probably prefer to be told something new. Since the new information provided by these experiments lies primarily in their implications, it will be helpful to make these implications explicit.
Of the skills we discussed, the one most directly relevant to comprehension instruction is main idea identification. The experiments showed that regardless of age, children have better memory for important than unimportant information in a passage. As we noted, however, recall does not necessarily reflect an ability to identify main ideas. The best way to find out if children can identify main ideas is to ask them directly, ideally with the text available to minimize memory demands. Although Brown and Smiley's (1977) importance ratings were obtained in such a way, the task complexity undoubtedly led to a low estimate of children's abilities. Using a much simpler task, Danner (1976) found that second graders could identify main ideas with some success. However, the passages he used were so short and simple that the older children may have found them insultingly easy. (This problem can arise whenever there is a large age range among subjects; materials that are the right level of complexity for one age group may not be appropriate for another.) Thus, we do not have much data on older children's main idea identification skills with more challenging passages. Moreover we do not know how skill at identifying main ideas changes with age.

We do know that there are developmental differences in the types of information children judge to be most important in stories (Stein & Glenn, 1978). There are undoubtedly individual and cultural differences as well, since everyone comes to the reading situation with different background experiences. However, the nature of the educational process requires that such differences be ironed out, for students are expected to extract the main ideas from their textbooks. Just how children learn to identify this normatively important information remains to be investigated.
The research on children's understanding of logical structure has fewer parallels in educational practice than the main idea research. Although children are often asked to unscramble pictures or sentences to create a logical story, there is little attention devoted to teaching how and why ideas within a passage are interconnected.

One reason for this lack of direct instruction may be that teachers feel children already understand logical relationships by the time they start to read. The research we reviewed demonstrates that preschoolers are, in fact, sensitive to logical structure in oral and picture narratives. The primary developmental difference in understanding logical structure seems to be in the strategies that are available for dealing with disorganized passages (Poulson et al., in press; Stein, 1976; Stein & Nezworski, 1977; Handler, 1978). Although we don't really know how these strategies develop, experience alone must be an important factor.

In view of the increasingly dominant role of expository texts in the child's educational experience, further research on understanding expository text organization is needed. Although Danner's contribution is important, additional studies should extend his work using more complex materials. Such research would be valuable not only for comprehension instruction, but also for instruction in writing; children must understand logical organization in order to write logically organized prose.

The research we reviewed on inferences provides us with little more than the conclusion that children can draw inferences when asked questions about sentence triplets and simple stories. The extent to which children
spontaneously draw inferences remains an empirical question, although there is some evidence that children will infer information that is crucial to comprehension (Brown et al., 1977; Stein & Glenn, 1978).

Given the methodological problems inherent in the inference research, the following generalizations should be regarded as tentative. There seems to be no evidence of developmental change in children's abilities to make inferences from sentence or picture triplets (Paris & Carter, 1973; Paris & Mahoney, 1974; Kail et al., 1977). However, there were developmental differences in making inferences from stories. Paris and Upton (1976) reported that older children were better at making contextual inferences than younger, and Stein and Glenn (1978) reported age differences in children's ability to infer missing elements in a story. These discrepant findings may result from differences in the scope of the required inferences. That is, inferences based on the sentence sets could be made by considering two simple sentences, whereas with stories, inferences often dealt with the theme of the story as a whole. Thus, younger children may have had difficulty considering the many components of a story simultaneously; this conclusion is supported by the fact that when the inferences dealt with specific words and phrases from a story, the developmental differences were eliminated (Brown et al., 1977; Paris & Upton, 1976).

An important comprehension skill that we did not touch upon in our review of the literature, but which is related to inferencing, is the ability to consider new material in light of what is already known. Little or no research has focused on this higher level aspect of comprehension (which some do not consider to be comprehension per se, rather applying the products
of comprehension), primarily because appropriate questions are difficult to formulate and are rather subjective. Nevertheless, this skill is crucial in answering "application" and "integration" questions frequently encountered on tests. Since even college students have difficulty with such questions, it is unlikely that young children consider incoming information with regard to a broader context of experience. By focusing too much on typical memory tests of learning and comprehension, such as free recall and recognition, educators run the risk of restricting student's intellectual creativity. Every teacher, for example, has probably encountered students who knew their course material by heart but failed a test because they were required to do come creative, integrative thinking. (See Baker & Santa, 1977, and Baker, Santa & Gentry, 1977, for empirical demonstrations of this phenomenon.)

Though the necessity for such "transsituational" comprehension increases as children become more involved in studying for content courses, it is probably never too early to introduce training in this skill.

The research dealing with children's understanding of story structure showed that children do in fact know what kinds of information belong in stories. Even four- and five-year-olds are capable of constructing well-formed stories that include purposive behavior. The research shows that young children have excellent comprehension of stories that conform to the structure specified by the schema. However, their comprehension is impaired when stories deviate from the schema, and this impairment is greater than that which occurs for older children and adults. One source of this developmental difference is less familiarity with discrepant structures. Clearly,
repeated experience with such stories allows one to build up strategies for dealing with them.

An important practical application of this research is to use the story grammar as a model for construction of instructional materials. Many of the materials currently prepared for beginning readers are sometimes little more than strings of sentences, lacking the conflicts and goals that are such crucial elements of stories (Bruce, 1978). It's no wonder, then, that many children regard reading as a boring task and not worth the effort. However, if stories were constructed to conform to a story schema, not only would the children find them more comprehensible, but hopefully they would discover that reading can be intrinsically rewarding.

While it is important for beginning readers to enjoy reading, it is also important that they learn to read expository prose, a task that is usually not nearly as much fun as reading a good story. Virtually all of the experiments on prose comprehension development have used stories as stimulus materials. One reason for this focus is to maintain children's interest in the task, but the primary advantage of using stories is that their structures can be specified by story grammar analysis. Nevertheless, researchers must also investigate expository comprehension, particularly in children of the "transitional" period; i.e., third and fourth graders who have mastered basic decoding skills but are not yet fluent readers. It is often at this time that reading problems become apparent, both because of the shift in emphasis from decoding to comprehension and because the children are expected to deal with expository prose in their social studies and science books for the first time.
Although there are undoubtedly many commonalities underlying comprehension of stories and expositions, there are also many differences which should be explored. We know that children understand stories quite well at an early age, yet we do not know how well they understand expository prose. It is possible we would find something akin to what Piaget (1952) has termed a "horizontal décalage": a particular child may be quite capable of performing a certain mental operation (i.e., making an inference) with a story but not with an expository text. Similarly, as we suggested earlier, children may understand logical structure in narrative before expository text. A number of factors may contribute to this hypothesized décalage, the most obvious of which is the child's greater experience with stories. In addition, stories have a higher order structure specified by cultural conventions, while expository text structures are more variable and ill-defined. Thus, children can use their story schema to enhance their story comprehension; no such generic knowledge is available for expository prose. Finally, stories are more concrete, with events and characters that the child can identify with, through experience or imagination. Expository material, on the other hand, is typically abstract, dealing with unfamiliar concepts and situations. In summary, since understanding is highly dependent on prior knowledge and experience, we should expect to find that young children have better comprehension of narrative than expository prose.

Our conclusion will conclude with a brief introduction to a new area of investigation, alluded to previously, that has important implications for educators: metacognition. Metacognition refers to the knowledge or awareness
people have about cognitive processes (e.g., memory, attention, comprehension, communication). A number of experiments have shown that young children are deficient in a variety of metacognitive skills (Brown, 1975, and in press; Flavell, 1978; Flavell & Wellman, 1977). For example, we noted previously that children seem to lack metacognitive knowledge about importance and logical organization (Brown & Smiley, 1977; Danner, 1976). But of more direct consequence to comprehension instruction is the growing evidence that young children have poor "metacomprehension" skills; that is, they do not always know when they don't understand. A recent study by Markman (1977) provides a good demonstration of this phenomenon. Children in grades 1-3 were given instructions on how to play a game or perform a magic trick. In both cases, information was left out that was critical to being able to follow the instructions. After listening to the instructions, the children were asked a series of questions designed to get them to indicate that they didn't understand. The children were told that their help was needed in coming up with good instructions, and that they should let the experimenter know if something was omitted or wasn't clear.

The older children asked questions much more readily than the younger, realizing that the instructions were incomplete. It was often not until the first graders actually tried to carry out the instructions that they realized they didn't understand. Markman concluded that this metacomprehension failure occurred because first graders did not execute the instructions mentally as they listened to them. Although their passive listening may have given them a feeling of understanding, because they didn't actively evaluate whether the
instructions made sense, they didn't know they didn't understand. In this experiment, then, children as young as third grade showed good metacomprehension. However, when the task demands are more complex, even college students are not very good at monitoring their comprehension (Baker, 1978).

These data suggest that keeping track of the state of one's comprehension during reading may be crucial to comprehension. This implies that poor comprehenders may benefit from metacomprehension training. Furthermore, it suggests that rather than wait until remediation is necessary, efforts should be made to teach metacomprehension skills in parallel with comprehension skills. At present, it seems that teachers do much of the metacognitive work for children (Wertsch, in press); the burden should be shifted to the children themselves. Further research should reveal that increasing children's awareness of their ongoing comprehension processes will enhance their comprehension skills.
References


Baker, L. Knowing when we don't understand (and what to do about it). In P. D. Pearson (Chair), What's new in reading? Clinic presented at the meeting of the American Association of School Administrators, Minneapolis, July, 1978.


Prose Comprehension Skills


Reder, L. *Prose comprehension: A literature review*. (companion IRA monograph on adults)


Acknowledgements

The preparation of this chapter was supported by the National Institute of Education under Contract No. US-NIE-C-400-76-0116. We would like to thank Susan Goldman, Glenn Kleiman, Carol Santa, John Santa, and Tom Trabasso for their helpful comments on earlier versions of this chapter.
Once there was a big grey fish named Albert who lived in a big icy pond near the edge of a forest. One day, Albert was swimming around the pond when he spotted a big juicy worm on top of the water. Albert knew how delicious worms tasted and wanted to eat that one for his dinner. So he swam very close to the worm and bit into him. Suddenly, Albert was pulled through the water into a boat. He had been caught by a fisherman. Albert felt sad and wished he had been more careful.

Table 1
Category Breakdown of a Well-formed Story

| Setting | 1. Once there was a big grey fish named Albert who lived in a big icy pond near the edge of a forest |
| Setting | 2. |
| Initiating Event | 3. One day, Albert was swimming around the pond when he spotted a big juicy worm on top of the water |
| Initiating Event | 4. |
| Internal Response | 5. Albert knew how delicious worms tasted |
| Internal Response | 6. and wanted to eat that one for his dinner |
| Attempt | 7. So he swam very close to the worm |
| Attempt | 8. and bit into him |
| Consequence | 9. Suddenly, Albert was pulled through the water into a boat |
| Consequence | 10. He had been caught by a fisherman |
| Reaction | 11. Albert felt sad |
| Reaction | 12. and wished he had been more careful |
No. 1: Durkin, D. Comprehension Instruction--Where Are You?, October 1977. (ERIC Document Reproduction Service No. ED 146 566, 14p., HC-$1.67, MF-$0.83)

No. 2: Asher, S. R. Sex Differences in Reading Achievement, October 1977. (ERIC Document Reproduction Service No. ED 146 567, 30p., HC-$2.06, MF-$0.83)


No. 4: Jenkins, J. R., & Pany, D. Teaching Reading Comprehension in the Middle Grades, January 1978.


* No. 4: Alessi, S. M., Anderson, T. H., & Biddle, W. B. Hardware and Software Considerations in Computer Based Course Management, November 1975. (ERIC Document Reproduction Service No. ED 134 928, 21p., HC-$1.67, MF-$0.83)

* No. 5: Schallert, D. L. Improving Memory for Prose: The Relationship Between Depth of Processing and Context, November 1975. (ERIC Document Reproduction Service No. ED 134 929, 37p., HC-$2.06, MF-$0.83)


* No. 8: Mason, J. M. Questioning the Notion of Independent Processing Stages in Reading, February 1976. (Journal of Educational Psychology, 1977, 69, 288-297)

* No. 9: Siegel, M. A. Teacher Behaviors and Curriculum Packages: Implications for Research and Education, April 1976. (ERIC Document Reproduction Service No. ED 134 932, 42p., HC-$2.06, MF-$0.83)


No. 16: Jenkins, J. R., & Pany, D. Curriculum Biases in Reading Achievement Tests, November 1976. (ERIC Document Reproduction Service No. ED 134 938, 24p., HC-$1.67, MF-$0.83)


No. 20: Kleiman, G. M. The Effect of Previous Context on Reading Individual Words, February 1977. (ERIC Document Reproduction Service No. ED 134 941, 76p., HC-$4.67, MF-$0.83)


No. 28: Ortony, A. Remembering and Understanding Jabberwocky and Small-Talk, March 1977. (ERIC Document Reproduction Service No. ED 137 753, 36p., HC-$2.06, MF-$0.83)


No. 30: Goetz, E. T., & Osborn, J. Procedures for Sampling Texts and Tasks in Kindergarten through Eighth Grade, April 1977. (ERIC Document Reproduction Service No. ED 146 565, 80p., HC-$4.67, MF-$0.83)


No. 34: Bruce, B. C. Plans and Social Actions, April 1977. (ERIC Document Reproduction Service No. ED 149 328, 45p., HC-$2.06, MF-$0.83)


No. 54: Fleischer, L. S., & Jenkins, J. R. Effects of Contextualized and De-contextualized Practice Conditions on Word Recognition, July 1977. (ERIC Document Reproduction Service No. ED 144 043, 37p., HC-$2.06, MF-$0.83)


No. 59: Mason, J. M. Reading Readiness: A Definition and Skills Hierarchy from Preschoolers' Developing Conceptions of Print, September 1977.

No. 60: Spiro, R. J., & Esposito, J. J. Superficial Processing of Explicit Inferences in Text, December 1977. (ERIC Document Reproduction Service No. ED 150 545, 27p., HC-$2.06, MF-$0.83)

No. 61: Brewer, W. F. Memory for the Pragmatic Implications of Sentences, October 1977. (ERIC Document Reproduction Service No. ED 146 564, 27p., HC-$2.06, MF-$0.83)


No. 77: Nash-Webber, B. L. Inference in an Approach to Discourse Anaphora, January 1978. (ERIC Document Reproduction Service No. ED 150 552, 30p., HC-$2.06, MF-$0.83)

No. 78: Gentner, D. On Relational Meaning: The Acquisition of Verb Meaning, December 1977. (ERIC Document Reproduction Service No. ED 149 325, 46p., HC-$2.06, MF-$0.83)


No. 81: Shoben, E. J. Choosing a Model of Sentence Picture Comparisons: A Reply to Catlin and Jones, February 1978. (ERIC Document Reproduction Service No. ED 150 577, 30p., HC-$2.06, MF-$0.83)


No. 83: Reynolds, R. E., Standiford, S. N., & Anderson, R. C. Distribution of Reading Time when Questions are Asked about a Restricted Category of Text Information, April 1978.


No. 95: Reichman, R. Conversational Coherency, July 1978.


No. 98: Green, G. M. Discourse Functions of Inversion Construction, July 1978.


No. 100: Jenkins, J. R., Pany, D., & Schreck, J. Vocabulary and Reading Comprehension: Instructional Effects, August 1978.
