This report describes the rationale and procedures for a series of studies related to reading tasks and reading texts. Children's reading tasks are examined in five strands of investigation: ability to read for meaning, ability to read for doing, ability to read for remembering, comprehension difficulties in skilled readers, and classroom observations. The text analyses consider six types of variables: orthographic-level variables, lexical-level variables, sentence-level variables, discourse-level variables, content-level variables, and other, extralinguistic variables. Overall, the report assumes that reading is a process of imposing an interpretation on a piece of textual material and that comprehension problems arise whenever an author's assumptions about the knowledge, skills, processing capacity, and expectations of the reader are unfulfilled. (AA)
CENTER FOR THE STUDY OF READING

Technical Report No. 43

THE ANALYSIS OF READING TASKS AND TEXTS

Allan M. Collins
Bolt Beranek and Newman Inc.
Ann L. Brown, Jerry L. Morgan,
and William F. Brewer
University of Illinois at Urbana-Champaign

April 1977

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

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

The research reported herein was supported in part by the National Institute of Education under Contract No. MS-NIE-C-400-76-0116. The assistance of Thomas H. Anderson, Steven R. Asher, Ernest T. Goetz, Georgia M. Green, Diane L. Schallert and Manfred B. Sellner is gratefully acknowledged. David E. Rumelhart and Andrew Ortony served as editors.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Part One - Task Analysis</td>
<td>5</td>
</tr>
<tr>
<td>Part Two - Text Analysis</td>
<td>32</td>
</tr>
<tr>
<td>Summary</td>
<td>77</td>
</tr>
</tbody>
</table>
Introduction

Too often educators and psychologists draw conclusions about research on an educationally relevant problem without adequate analysis of the nature of the problem itself. We feel that it is important to begin our study of reading comprehension with a detailed account of the tasks the young reader faces. We feel it is important here to analyze both the details of the tasks themselves and the nature of the additional social, communicative, and related demands imposed by the classroom situation. Moreover we suspect that there are general teachable strategies that young children normally learn by trial and error which greatly facilitate his performance on these tasks. We therefore propose a study of these strategies with an eye toward making them an explicit part of reading education.

A central part of the tasks under analysis is, of course, the textual material with which the child is required to interact. Although an adequate analysis of a text cannot be accomplished outside of the context in which the reader interacts with the text, it is nevertheless useful to look very carefully at the nature of the particular texts children encounter at different stages of their educational development. At first the texts are primarily designed as vehicles for teaching the child decoding skills. By the fourth grade, however, the texts increasingly are intended to serve as vehicles for teaching content. Authors of these texts often attempt to control various aspects of the texts depending on the reason for which the texts have been devised. Texts devised for beginning readers in one way or another pay more attention to such things as orthographic structure and word frequency (see report on skills hierarchies). As the attitudes of the authors of textbooks vary, aspects of the text not under specific
control by the author often undergo greater variation. Moreover, there are aspects of a text (such as overall structural characteristics) which are rarely carefully controlled. It is important for the researcher, as well as those involved in curriculum development, to understand both what the important characteristics of currently employed texts are and the effects these characteristics might have on their comprehensibility for children at various stages of their development and with various cultural experiences. The accomplishment of this task requires: (1) that we develop methods of analyzing textual material which are applicable at all levels of analysis (orthographic, lexical, syntactic, semantic, aspects of textual pragmatic and general contextual/processing) which can serve to analyze all relevant genres of texts (instructions, stories, essays) and which can be used as a basis for the analysis of these texts for readers at all grade levels K-8; (2) that we investigate the effects on comprehension of these various aspects of the text on readers at different levels of reading ability, with different interests and with different socio-economic backgrounds. These are clearly long term tasks. It is doubtful whether adequate answers to these questions can be developed even within several years. Yet we believe that a frontal attack on them is essential to the mission of the Reading Center. We are, of course, not beginning in a vacuum. Fortunately, recent trends on textual analysis in education, psychology, linguistics, and artificial intelligence, all work in our favor and will contribute substantially to our efforts.

In what follows, we proceed by first laying out our current research plan for the analysis of reading-relevant tasks. In this section we give special attention to the various reasons for which children read—reading
for meaning, for doing, or for remembering. We look at the strategies children use and can be taught to use in the abstraction of main ideas from their texts. We propose to study the classroom—how do teachers and children really interact in the teaching of reading skills? We propose to look at the kinds of inferences that are demanded of children when they read texts designed for them, at their ability to make these inferences, and at what they do when they cannot make them. We propose to look at the reading strategies employed by skilled readers when trying to construct interpretations of difficult to understand materials. Perhaps such strategies can be taught directly as part of reading education. We then turn to a discussion of our current thoughts on text analysis.

Texts, of course, can differ from one another on many dimensions. An adequate theory of textual processing must account for the ways textual and situational characteristics at these different levels can interact to give a skilled reader a coherent interpretation of the text. We must determine just what aspects of the text readers at various levels can attend to. In the section on textual analysis we describe existing and proposed research into the textual characteristics relevant to the orthographic, lexical, sentential, paragraph, and discourse levels. At each of these levels our research plans are to focus on characteristics of existing texts and how these characteristics interact with readers with different knowledge bases reading for different purposes.
Part One - Task Analysis

I. Introduction

II. Children's ability to read for meaning

III. Children's ability to read for doing

IV. Children's ability to read for remembering

V. Determination of comprehension difficulties in skilled readers

VI. Observation in classrooms

VII. Conclusion
Introduction

The goal of developing task analysis procedures is to make in-depth analyses of different aspects of reading in order to pinpoint where problems in comprehension arise during the transition period. Such an analysis can tell us both what reading skills should be taught and how to diagnose where children's difficulties arise. This research should look at children and adults performing typical tasks, using a variety of techniques to analyze their capabilities and their underlying processing.

A main focus of this section is on how one goes about studying reading comprehension and study skills in the less-mature reader. The main problem is one of externalizing an internal mental event; the younger the subject the less likely it is that direct questioning will result in any worthwhile information. Not only is the younger child less able to express himself, but he is also less aware of his own cognitive processes and less familiar with the self-interrogation techniques necessary to achieve the degree of introspection needed for analyses such as those possible with adult readers. Thus it is important to develop methods of eliciting evidence of the child's self-knowledge, or lack of it, other than by means of direct questioning.

As a second general point, it should be noted that we know very little about the skills possessed by mature readers, let alone about the developmental sequence in which they are acquired. The development of a descriptive theory of the skilled reader's strategic processes is a prerequisite for informed developmental research. Only when we understand basic commonalities in the reading strategies of the efficient can we look intelligently for problems which can be the subject of remediation.
It is not difficult to predict that we will find that older children are more strategic than younger children, and know more about reading effectively. It will also fail to amaze us if we find disadvantaged children particularly ignorant in this field. What is necessary is insight (based on our insights of mature processing), of particularly prevalent pockets of inefficiency which are trans-situational, i.e., a strategy deficit which could effect performance on a wide variety of reading tasks. The identification of such deficits and the planning of suitable training programs is, of course, the aim of instructional psychology. In order to concentrate efforts on high pay-off areas, we need the information concerning what are the essential strategies the adult possesses.

Another general point which should be considered in developing task analysis procedures is a taxonomy of reading purposes. A taxonomy suggested by Sticht (personal communication) is increasingly useful to us in thinking of what to look at in children.

(a) **Reading for meaning** - reading in sufficient depth that one can extract the gist, with no intent to meet future tests; pleasure reading would fall into this category.

(b) **Reading for doing** - reading instructions, signs, billboards, bus timetables, recipes, driving instructions, etc.

(c) **Reading for remembering** - deliberate studying in anticipation of a future test.

These different tasks demand different levels of processing, different skills and strategies, and different levels of awareness. For these reasons, they should be kept somewhat separate. In addition, different subject populations need varying levels of skills on the three. For example, educable retarded children, and many others who do not, or can not, enter
academic pursuits, do not need to be efficient in studying. They rarely have to face examinations in their job situation. Training aimed at this population might be more profitable if it were focused more on reading for doing. This is the kind of reading that is required in everyday life situations, - one must read signs and notices, and job manuals and recipes, etc. Other types of reading are a luxury. In the "Cloak of Competence" (Edgerton's, 1967 description of retarded individuals living "normal lives" in the community), only two of the 110 members of the cohort were found to read for pleasure (novels, etc.), with the exception of illustrated papers like comics and Daily News. Most of the cohorts, on the other hand, experienced reading problems of the reading for doing variety, and were troubled by them. They needed to read advertisements of grocery specials, to follow instructions on assembling products. Intervention and training programs aimed at instruction-following seem particularly necessary for such a population and would certainly have life-improving consequences if successful.

With respect to these different types of reading, there are a number of questions that should be analyzed with task analysis procedures:

1. What comprehension skills do skilled readers acquire for dealing with different reading tasks?

2. Which of these skills do different children acquire or fail to acquire?

3. When do different reading skills develop from kindergarten to adulthood?

4. How well does what is taught in schools teach the various skills necessary for skilled reading?
5. As reading tasks change, how do the skills required change?

6. How well can the various skills that are needed for skilled reading be taught directly?

Because these different questions require different techniques, this section proposes a number of different procedures for performing task analysis. Such diversity permits the selection of the most successful techniques, and ensures that the task analyses will converge on the most important problems that arise in reading comprehension. The section describes a number of different task analysis procedures that are currently being piloted or considered as potential techniques that should be applied.

Children's Ability to Read for Meaning

A fundamental skill for efficient reading in school is the ability to abstract the gist, or main ideas, from a text. Our work on this topic with children is in the preliminary stages and has largely centered on methods of obtaining measurable indices of the child's knowledge. To date we have completed initial studies with normal children of from kindergarten to eighth grade, and retarded children of approximately two years delay for each age. In general, we know that young children are better able to pick out the main idea of picture sequences than of stories presented verbally. They are also more likely than older children to be misled by red herrings, i.e., an irrelevant part of the picture made physically larger, or an irrelevant part of the story repeated or emphasized.

In examining children's ability to study differentially the main ideas, we have attempted to obtain scoreable data by having the child underline
key words or sentences, underline or circle key words or write notes. To summarize we find:

(a) Children below seventh grade, and junior high school educable retarded children do not tend to introduce such activities of their own volition.

(b) Even when directly instructed to do so, they do not use the techniques intelligently. For example, children who underline or circle words spontaneously tend to underline or circle words or idea units previously rated by adults to be key or important elements. Children who do not use the technique spontaneously, but only when instructed, tend not to underline key words but appear to behave randomly; or, they circle or underline long or unfamiliar words. We even have examples of children told to underline important points who underline everything! If they are told to underline, they will, but not strategically as spontaneous users tend to do.

These studies are in their infancy, pilot data is not yet completed and technique refinement is still underway. Of importance is that these data are readily scoreable. For each child we have a marked study sheet. We can compare the underlined or circled words against adults' scoring of key or important elements, as in the Brown and Smiley (1977) study.

Again tentatively, it looks like there will be a hierarchy of these skills; circling key words is easier than underlining key idea units, and note taking is particularly late in developing.

We plan to systematically observe the development of note-taking competence in the junior high and high school years. Specifically, we will be looking for:

(a) spontaneous use - when do children first start to take notes
(b) **quality of notes** - this can be evaluated by having adults rate the clarity, organization, and representative nature of the child's productions.

(c) **use of note-taking** - more difficult to assess - does the child take notes to solidify and organize what he is reading, or to provide a record which consulted, or both. Does note-taking help performance, etc.

Several aspects of story understanding are important in our attempts to investigate children's abilities to read for meaning. We will discuss story recall data as a means of quantifying this ability, and an analysis of story understanding as a means for investigating higher level aspects of reading for meaning.

In recounting stories it seems that even young children favor the main ideas. However, they tend to recount simple action sequences, and rarely mention internal states (Mandler and Johnson, 1977) or moods.

Although children remember the main message even in the absence of a deliberate attempt to concentrate on essential ideas, this does not exclude the possibility that deliberate strategies could be used to enhance comprehension and recall of prose materials; indeed, it would seem unreasonable to suppose that the mature individual would not possess a rich variety of such skills. We suspect that the mature learner's awareness of his skills and of the task at hand would be particularly useful if longer preparatory periods were permitted for the study of material which must be maintained over some period of time, a more typical school learning situation. For example, in the Brown and Smiley (1977) paradigm, we would predict that
older children would use their knowledge of the importance of certain pieces of text to study efficiently, thereby enhancing their recall of important material. Younger children, not possessing the necessary "meta-comprehension" skills, would not benefit as much from the increased study time to focus on the essential. Thus their recall would improve little if at all. We propose to investigate this hypothesis.

The situation changes, however, if one trains children to use self-testing mechanisms for recall readiness. We have been working on this for some time. The basic idea is that the child is told to study material in anticipation of a test. He is to indicate his readiness for that test. He has unlimited time. The measures of effective recall readiness are: time taken before test, amount recalled at test, and observed activities during the study period. This is a very simple paradigm and can be used for simple list learning rote recall situations (Brown & Barclay, 1976) and learning from prose. Even retarded children can readily be trained to wait before a test while they undertake some rudimentary self-monitoring procedures - and then they do benefit from the delay. Training on a list recall task generalizes to prose recall (Brown & Campione, 1977). We intend to extend this paradigm to more realistic study situations, and observe and interrogate children in the study period.

If the ability to rate units of a complex passage in terms of importance to the theme is a late developing skill, the implication for training study habits, reading comprehension skills and even the design of text materials are important. Children who have difficulty determining what are the key points of a passage, can hardly be expected to select them for intensive study. Of interest in this context is an experiment by
Hershberger and Terry (1965) concerned with text learning in eighth-grade students. One variation which improved performance was printing all core or essential contents in red and all non-essential material in black. That is, some form of specific training or cueing in identifying important features of a passage may be needed before efficient study behavior will form a consistent part of the child's repertoire of cognitive skills. The task of identifying (and even training) such skills, particularly in slow learning children, will be a major feature of our work in this area. Mature readers appear to engage in self-interrogation to monitor whether they are comprehending the important points in a text. We propose to (1) provide good questions for young readers to model the kinds of questions they should ask themselves, (2) withdraw the questions and instruct children to generate questions of the type modeled, and (3) determine whether this can be made a self-sustaining study strategy.

Another major aspect of reading comprehension that should be studied with task analysis is children's ability to make inferences about the motivations and actions in a story, and about the author's intentions in writing a story. The technique we have developed involves questioning children about stories given them with respect to a formal process theory of reading comprehension (Adams & Collins, in press; Rumelhart & Ortony 1976). This technique is an extension of techniques developed for the analysis of human inference strategies in answering questions and human tutoring strategies (Collins, Wainock & Fassafiume, 1975). The goal of this work is to determine what inference skills children develop at different points in the transition period, and whether they have developed
the necessary skills for comprehending the texts that are encountered during the transition period.

For our initial investigation in this area we have collected four short stories that are quite difficult to understand in terms of the motivations of the actors and the intentions of the authors. The stories include three fables from Aesop and a story about Ben Franklin from the Education Testing Service (1960) Cooperative English Test of reading comprehension. These stories have been rewritten to make the vocabulary and sentence structure comprehensible to kindergarten children, and thus to remove these variables as sources of difficulty in comprehension.

The procedure we will use is to have some children read the stories. We will encourage them to read for understanding and to reread whatever they do not understand. When they think they understand the story, they will be asked to retell it in their own words what happened, and what they thought the point of the story was. Then they will be asked a series of specific questions about each story. The questions are of four types: 1) about the world knowledge necessary to understand the story, 2) about the information in the text, 3) about the inferences necessary to understand the various actions and motivations in the story, and 4) about the author’s intentions in writing the text. (One of the stories, and the questions about the story are shown in Tables 1 and 2.) This kind of questioning often leads the reader to further interpretations of the story, but the questions are ordered to minimize such effects. Even if children do revise their interpretations, the comprehension failures isolated by systematic questioning are likely to help us pinpoint where comprehension difficulties arise in the reading process.
A poor man came to a large house during a storm to beg for food. He was sent away with angry words, but he went back and asked, "May I at least dry my clothes by the fire, because I am wet from the rain?" The maid thought this would not cost anything, so she let him come in.

Inside he told the cook that if she would give him a pan, and let him fill it with water, he would make some stone soup. This was a new dish to the cook, so she agreed to let him make it.

The man then got a stone from the road and put it in the pan. The cook gave him some salt, peas, mint, and all the scraps of meat that she could spare to throw in.

Thus the poor man made a delicious stone soup and the cook said, "Well done! You have made a wonderful soup out of practically nothing."
TABLE 2
Questions on Stone Soup

1. Can you tell me what happened in the story?
2. What was the story meant to show?
3. Was there anything you didn't understand when you first read it?
4. Why did the man suggest making stone soup?
5. Why did the cook think the man had done something clever?
6. Why did the soup taste so good?
7. Why did the man ask to dry himself by the fire?
8. Why did the maid send the man away at first?
9. Why did the maid finally let the man come in?
10. Why did the cook let the man make stone soup?
11. What did the man come to the house for originally?
12. Did the stone make the soup taste good?
13. What else was in the soup?
14. What do you think of the poor man?
15. How do you make home made soup?
16. Did the man trick the cook in any way?
17. Will a stone change in any way if it is heated in water?
18. Can a stone be eaten?
19. Was the story meant to show that stones are good for making soup?
20. Was the story meant to show that if you want something enough there is always a way to get it?
21. Was the story meant to show that it is possible to make something good, out of bits and pieces?
This study illustrates one kind of task analysis procedure that can be applied to a variety of reading tasks that children of different ages are given. The procedures described should be used for those tasks that the survey shows are representative of reading tasks children face in school. One kind of reading task where the inferential problems are often of similar magnitude to reading fables is that of reading instruction. Instructions are often underspecified, so that children are faced with the problem of figuring out what are possible meanings of the instruction, given the constraints the real world imposes on the task (e.g. assembly instructions are constrained by the nature of the components available). But systematic questioning is only one technique that can be used to study inferential difficulties. Other task analysis procedures (described in this section) should also be used to study the children’s capabilities in performing different reading tasks.

**Children’s Ability to Read for Doing**

Task analysis of how children read instructions should have very high payoff (a) because of its importance for everyday life, particularly among retarded children, and (b) because it is easier to externalize the child’s comprehension failures on a doing task -- if he does not understand he will not perform some action correctly. Thus, the experimenter will be signalled that a comprehension failure has occurred and can intervene, question the child, analyze the error patterns, etc.

A good way to examine monitoring, detection and recovery mechanisms, is to set up situations where the child must read the steps in a set of instructions and then perform a task. If the task can be physically separated from the instructions this will further serve to externalize the child’s
sequence of actions. The set up would be as follows: The child is given a sequence of instructions such as those of a recipe. The recipe will be in a central location, e.g., in a recipe card holder. The child must read the steps and then perform actions at a geographically-separated location, e.g., read at the desk and perform at a stove or cutting board. In a kitchen situation such separation is by no means atypical. The separation is important as the child must read and then move away to perform. Presumably, moving away from the instruction would indicate that he is ready to perform the next step. Checking back would also require physically returning to the instruction; again, this could be easily noted by the observer, and questions could be asked of the child concerning his understanding of each step and his checking back behavior.

Recipes do seem particularly suited for the task, as they require checking and monitoring as well as some commonsense assumptions. Consider this example from Crock Cookery by Mike Roy.

French Onion Soup

4 lbs. butter  
6 onions  
2 cans beef broth  
1 1/2 soup cans water  
6 slices French bread  
3/4 cup grated Parmesan cheese

In a skillet, melt butter and cook onions until lightly browned. Put them into the crock cooker along with the remaining ingredients. Cook on low for 4-6 hours. At serving time, spoon soup into heat-proof cups and top with toasted bread. Sprinkle with cheese. If convenient, place cups under broiler until cheese bubbles.
This book is being considered for our pilot studies on following recipes because the recipes are simple and the words easy to understand. But note the commonsense checking problem involved. If you follow the instructions in the exact sequence, without reading the recipe through first, you put the bread and cheese into the soup and cook four hours. Result: disgusting soggy bread in soup, and no crisp bread and cheese to top the soup with! The demand for commonsense inferences like this is rampant in cookbooks, and particularly in so-called simple ones, where only a few steps are mentioned.

Assembling toys is another promising area we are looking at. The task seems to have the same properties and can be treated like the recipe-following one. With educable retarded adolescents, we feel it necessary to choose material which they spontaneously try to read. From considering Edgerton's ethological descriptions of real-life activities of retardates in the community, we learned that they are really interested in driver education manuals. Teachers at the high schools confirm that the children are extremely keen to learn to drive. Thus such manuals seem particularly suited for studies with these populations.

Children's Ability to Read for Remembering

In considering the way in which children read for the purpose of remembering what they have read, it is clear that the ability to grasp the main ideas is important. We saw earlier that underlining and note-taking are useful strategies toward this end, as is self interrogation with respect to recall readiness. In this section, we review other study strategies that children learn and employ, and we discuss ways in which some of them can be trained.
We are currently engaged in an exploratory investigation of students' study habits, or how children think they go about studying. We intend to conduct a semi-structured interview. We are devising a series of set questions that will be put to the students in the context of an information session where they will be encouraged to talk freely about exactly what they do when faced with various forms of study assignments (time spent, reading strategy, study strategy). The interviewer will practice various probe questions to elicit more information if the students are reticent.

The aim of the project is to obtain preliminary data concerning how children set about studying. We are particularly interested in comparison between the seventh-grade students and seniors who may have acquired idiosyncratic study routines. The senior class responses will be compared with those obtained from college students.

On the basis of this information, we hope to develop a formal interview technique to use as a diagnostic tool with the students from other schools particularly those who are less successful at academic pursuits. This would be a first step in our intended program for developing training procedures.

Because we anticipate difficulty in obtaining adequate verbal reports and introspections from younger children, we hope to use a cross-age tutoring program as a vehicle for forcing the children to explicitly state what they believe to be an efficient study plan. Children will be enrolled in a program of tutoring at children (a) close to their own age, (b) much younger. They will be asked to help teach these less efficient learners to study effectively. In devising these instructional plans they will have the advice of an experimenter-observer. By observing their questions and plans concerning what to teach, we hope to gain valuable insights into
their knowledge concerning "how to study." Adequate debriefing for poorly instructed trainees will be provided.

A similar idea is the little teacher idea -- the aim is the same, to provide a "real-life" situation which might elicit from the child any information he has concerning reading strategies. The basic idea is that children will be divided into small groups to participate in an academic project, e.g., a science demonstration. One child will be selected as the instructor each session. It will be his job to read the material, provide a summary for the others, and provide illustrations and demonstrations for his confederates. In the preparation he will be aided by the experimenter-observer as in the cross-age tutoring project. Each team will be told that they will be tested on the material and they must try to be the best team (to win a prize, etc.). The child's knowledge or ignorance concerning the main ideas, emphasis, illustrations, etc., and how to study material should be made explicit by this procedure.

As an extension of previous work with video-tape models (Brown, Barclay & Jones, 1977), we are constructing video-tapes of children performing various study activities, such as rereading, underlining, self-testing, note-taking. We will see if observing children will model their performance after the tapes. Of more importance we will use the observation period to elicit evidence of the child's own knowledge concerning reading strategies -- e.g., ask the observer why the model is undertaking a certain activity, if it will help him, does the observer do that, etc.

As an extension of Markman's (1977) work on metacognition, we will ask children to act as critics. Specifically they will be asked to rate the adequacy of a set of instructions we have written, e.g., a book
of recipes, or a manual on how to use a toy, or play a game. The instructions will be obviously incomplete, or misleading as in the onion soup example. Children will be asked to help us rewrite faulty instructions so that children of their age can read them. Again when informing us of the inadequacies of our prose, they should be revealing their comprehension of reading, and instruction-following.

**Determination of Comprehension Difficulties in Skilled Readers**

Comprehension difficulties often arise in the transition period because children move from reading texts designed to teach basic reading skills to texts that serve some educational purpose, such as expository texts and instructions. The early texts are based on vocabulary and situations that are relatively familiar to children. But the latter kind of texts involve new ideas, new vocabulary, etc. Thus, a new set of problems arises such as how to deal with failures to understand different words and phrases, how to revise misinterpretations, how to select the important points for whatever purpose is at hand (e.g., doing the task if given instructions, or remembering later if given expository material). These require a set of skills that children have not needed before. To some extent the same problems arise in aural contexts, but the appropriate strategies for dealing with such problems are quite different in aural and reading contexts. Hence, children have not learned strategies for dealing with such problems, and for the most part, they are not taught them. In such situations the brighter children develop strategies for coping and the less bright children lose interest.
An important goal of task analysis therefore, should be to determine what strategies skilled readers have developed for dealing with such problems. Then further task analysis should identify which of those strategies children learn, when they learn them, which children learn them, and whether children can be taught them.

We can illustrate the importance of this kind of task analysis in terms of our preliminary analysis of skilled reader's strategies for dealing with comprehension failures. Coping with school must depend heavily on being able to figure out how to make sense of texts when the meaning is not obvious to the reader because of insufficient knowledge. We have carried out a preliminary study in which adult readers were given difficult expository texts to read and comment on. After reading each paragraph, they were asked to comment on the following issues:

1. What sentences did you fail to understand completely?
2. What was the nature of the failure?
3. Did you have any hypotheses as to what the meaning might be?
4. What sentences either confirmed or disconfirmed previous hypotheses?
5. If you do something other than read on, what do you do?
6. If you jump somewhere else, why?
7. Any other comments on what you do as you read.

There were several striking aspects of the response protocols collected so far: 1) many words and phrases were only partially understood, 2) the reader would often formulate hypotheses about possible meanings, 3) later sentences often provided confirming or disconfirming evidence about previous hypotheses, and 4) the reader would sometimes jump backwards or forwards.
in the text to figure out the meaning. The subjects were mostly unaware of these strategic aspects of their reading beforehand; slowing down the process highlighted the complex strategies that they used.

These preliminary sessions led to a tentative theory of how skilled readers deal with comprehension failures.

1. If a reader reads something that he does not understand, he may decide to take some strategic action or store the failure in memory as a pending question.

2. If he stores it as a pending question, he may formulate a possible meaning (usually one) which is stored as a tentative hypothesis.

3. If he forms a pending question, he usually continues to read.

4. If a triggering event occurs after forming the pending question (e.g., too many pending questions or repetition of the same pending question) the reader may take some strategic action.

5. Possible strategic actions he may take are:
   a. reread some portion of the text to collect more information that will either answer a pending question or form a tentative hypothesis related to a pending question.
   b. jump ahead in the text to see if there are headings or paragraphs that refer to the pending question and which might answer the pending question.
c. consult an outside source (e.g., dictionary, glossary, encyclopedia, expert) to answer some pending question.
d. make a written record of a pending question.
e. reflect about information that the student has in memory related to the pending question.
f. decide to suppress a pending question because the text structure suggests it will be answered later.
g. quit reading the text.

6. If the strategic action is successful, the reader usually continues reading from the point he last encountered the comprehension failure.

7. If the strategic action is not successful, the reader usually takes some other strategic action.

The research questions raised by the model of study depicted in the previous section include the following: text characteristics affecting the reader's ability to form a pending question, the nature of pending questions, events which act to trigger strategic actions, types of strategic actions, and the decision process for selecting a type of strategic action.

Text characteristics affecting the reader's ability to form a pending question include the difficulty of vocabulary and logical structure of the text, necessary prerequisite information, and the depth of structure of the information being read. It is important to know the extent to which the nature of the text affects the student's ability to formulate pending questions, since it will ultimately affect the student's ability to do something when he does not understand what he is reading.
A more basic research question, however, concerns the nature of pending questions. That is, what is the form in which skilled readers go about keeping tabs on their understanding? The fact that we call it a "pending question", implies that it is indeed a question, but this might not in fact be so. The reader may ask and observe his ability to answer questions about the text, but he may also observe his ability to fill in missing information and details, to paraphrase the text, to apply concepts and principles being discussed, to produce mental images of the text content, or to follow the logical structure of arguments in the text.

Given that the pending question itself is not always the stimulus for taking a strategic action to correct misunderstandings, what are other stimuli which cause strategic action to be taken? Likely events which might act as triggering events include an overflow in the pending file (i.e., too many pending questions), the repetition of a pending question, the disconfirmation of a previously formed hypothesis (either a tentative or a highly confident hypothesis, although each might trigger different strategic actions), formation of a pending question which the text treats as "given" information (i.e., treating a math proof as "obvious" when the reader does not see it as obvious), encountering a pending question which the reader suspects is very important, or encountering new vocabulary words easily explained through recourse to a dictionary.

The next logical set of research questions concerns the types of strategic action which can be taken. Those indicated by the model described include
looking back or forward in the text, looking to an outside source (e.g., another text or a teacher), thinking and reflecting about the problem or confusion, formation of a tentative hypothesis about "what the text is trying to say", doing some sort of record keeping such as writing a note about the confusion, or quitting the study process for a while. We know that students engage in all of these activities on occasion, but which are the most appropriate for which kinds of understanding difficulties?

The last research question following from the model is the process by which the reader decides on (selects) a type of action. The major question here is whether there is any relation between the type of comprehension difficulty and the type of action. Another question is whether there is a default action which usually is resorted to if another is not obvious. For instance, rereading may be a common and/or sensible default strategic action if no others come to mind to the reader. If there is a relation between type of comprehension failure and type of strategic action, what is that relation? If there is none, what IS the selection of a strategic action related to?

If we can formulate a precise theory of how skilled readers deal with comprehension failures, then it will be possible to design experiments and perform task analyses to study how well children develop such skills. Further it would be possible to teach these strategies and study how well children acquire and maintain them. This same research approach can be applied to various aspects of skilled reading: picking out main points, revising misinterpretations, reviewing texts read earlier, and so forth.
The goal is to formulate process theories of the important skills adult readers have unconsciously acquired to cope with the reading demands made on them.

This proposed research involves a somewhat different use of task analysis than determining how children perform in various reading tasks (e.g., reading instruction or stories) they are given in the transition period. Instead, this kind of analysis, some of which is conducted under computer control, investigates how readers deal with specific aspects of reading, such as comprehension failures, that arise in a variety of reading tasks. Both types of analysis seem critical to understanding the comprehension difficulties that arise during the transition period.

Observation in the Classroom

Preliminary observation in an urban classroom indicated that there are many ways that the various reading tasks and exercises given children fail to accomplish the goals intended by curriculum designers. For example, the following kinds of problems arose when children were working exercises designed to teach vocabulary skills: 1) items were inappropriate to the level of skill or the background of the child, 2) tasks were of doubtful value to teaching vocabulary skills, 3) the children did not understand what to do, 4) the teachers did not know what kinds of problems children were having. Analyzing where such failures occur is critical to understanding why difficulties arise in reading comprehension during the transition period.
In order to analyze where difficulties arise, systematic observation should be carried out in classrooms covering the span from kindergarten to eighth grade. The kinds of questions such task analysis should address are the following:

1. If done properly, would the task actually teach the reading skills intended?
2. Does the teacher understand the nature and purpose of the task?
3. Does the teacher communicate the nature and purpose of the task?
4. Do the students understand the nature and purpose of the task?
5. Are the students motivated to do the task?
6. Do the students have the background and skills to do the task?
7. Do the students have the time and resources to do the task?
8. Do the students actually do the tasks?
9. Does the teacher know how well the students are doing and what difficulties they are having?
10. Does the teacher give help or feedback to the students, and is it important that s/he do so?
11. How important are the skills taught this way in the overall process of reading?

The task analysis procedures to be used should include tape recording and annotating events that occur in the classroom, and conducting interviews with teachers and students. In particular, the observer should record the teacher's instructions to the students and annotate what the individual student does (with respect to the above questions). Any further interactions between the student and teacher should also be recorded. A portion of the students
observed should be questioned while they are performing the task (in as unobtrusive and non-leading a manner as possible) to determine the reasons for whatever difficulties they may have. All students should be questioned subsequently about any difficulties they had, and what they thought they were supposed to do and why. Finally, teachers should be interviewed after class as to what they thought was the purpose of the task, how effective they thought it was, and how well they perceived what the students were doing and the problems they were having. Observations of this kind can only be carried out with a few students in a few classrooms, but it should be applied to every kind of reading related activity in the classroom.

These in-depth analyses of reading in the classroom are quite different from the sampling of task contexts in the classroom (Goetz & Osborn, 1977). The attempt here is to evaluate when and why comprehension difficulties arise in the classroom with respect to a theoretical analysis of the reading process. In this way we hope to discover which techniques now being used are most effective for developing student’s comprehension skills and which are least effective. By comparing what is done in the classroom to a theoretical analysis of reading, it suggests what new strategies for teaching reading comprehension should be attempted. Finally, by analyzing how what is done is failing, it may be possible to make recommendations that would render reading curricula more failure-proof.

Conclusions

We have outlined a rather substantial body of work to be carried out in the analysis of some of the actual tasks children must carry out and a
discussion of some of the strategies involved in doing so as children become mature readers. Of course, even this body of work will not give us a complete picture of the task demands imposed on beginning readers. We have, for example, not given enough attention to the interaction of these task demands with cultural differences and other individual differences in knowledge and attitudes toward reading. We have, however, laid out a program of research which will bring us a step closer to understanding reading "from the child's point of view." Such an understanding will not only allow us to pinpoint problems of poorer readers and successful strategies of good readers, but will also provide information for the ultimate modification of the task demands themselves bringing them more in line with the theoretical accounts of the reading process we are developing elsewhere.
Part Two - Text Analysis

I. Introduction

II. Orthographic level variables

III. Lexical level variables

IV. Sentence level variables

V. Discourse level variables

   Psychological approach

   Linguistic approach

VI. Content level variables

   Text structure and comprehension

   Descriptive analysis of text content

VII. Other extra-linguistic variables

   The role of illustrations

   The role of interest

VIII. Conclusion
Introduction

Clearly the nature of the texts that children encounter plays an important role in their ability to read and understand those texts. The problem on which we focus here is that of isolating and investigating the role and relative importance the relevant properties of texts possess, particularly as they pertain to reading. We are especially interested in those properties as they contribute to the "difficulty" of reading a text. An understanding of the role and importance of such properties can make it possible to establish reliable criteria for assessing the appropriateness of reading materials for various individuals, levels of development and for various tasks. This would allow, for example, the matching of text characteristics to the skills, knowledge, and interests of a particular student. It would also contribute to the design of textual materials both for testing levels of reading ability, and for teaching relevant reading skills.

Essential in this respect is the development of adequate theoretical models to serve as a basis for our analysis and to support the development of relevant descriptive techniques for classifying, evaluating, and coding texts. The development of such theoretical and descriptive tools must be our first order of business, for without these we can neither adequately describe the characteristics of current textual materials nor can we study how these variables interact with a child's state of knowledge, level of skills, and purpose in reading the text.

A central tenet of our theoretical perspective is that textual processing involves the creation of meaning by the reader based on the clues provided by the author in the text. These clues come in many guises.
When an author constructs a text he chooses to make certain kinds of marks on the page because he believes that his projected audience will interpret those marks according to certain conventions. Some of these conventions are very explicit and a good deal of reading education is focused upon these. Thus, for example, the conventions that certain configurations of marks will constitute the letter "a", or that the sequence of marks is to be interpreted from left to right and from top to bottom are well understood and knowingly adhered to. Other conventions, such as that certain configurations of letters correspond to certain patterns of sound are less reliable; in some cases authors are careful not to include too many words in which these conventions are violated. In other cases, however, words like "have", "would", "the", "love", etc., are used without realizing that they form special cases which violate this convention (although, of course, conforming to the conventional way of spelling these words).

The general linguistic and communicative conventions that are employed which are only tacitly understood by the author constitute a different kind of example. Here, although an author may well be able to competently use these devices in an appropriate way, he may not understand exactly what these devices are, and therefore may not appreciate the sophistication or cultural knowledge required for the appropriate use of them by the reader. For example, the proper comprehension of textual material invariably involves a rather large set of inferences that the author presupposes the reader is able to make. These presuppositions are often made without conscious awareness and therefore without conscious monitoring by the author. Thus, although textual material devised to be simple may well contain little new vocabulary or only simple syntax, an author may
inadvertantly make his material very difficult to comprehend because of all he has left unsaid.

The overall difficulty of a text is a complex function of the orthographic structure of the words employed, the difficulty of the vocabulary, the syntactic forms used at the sentential level, the organization of the sentences into higher level structures like paragraphs, the structural nature of the content of the text, the inferences necessary for comprehension, the role of illustrations, and so on. The role of each of these factors in a particular text obviously changes with the level of sophistication, knowledge, and motivation of the reader. It is also obvious that all of these levels interact in highly complex ways in the actual comprehension process, but for ease of exposition our proposed work will be broken down into the following somewhat arbitrary categories: of orthographic level, lexical level, sentential level, paragraph level, story level, and discourse level. We will begin with but a brief discussion of orthographic and lexical aspects of text structure and move rather directly to a discussion of the bulk of our proposed research on the sentential, paragraph, story, and discourse levels of analysis.

**Orthographic Level Variables**

Although orthographic complexity probably plays very little role in determining the overall difficulty of a text for mature readers, we presume that it is an important determinant of difficulty for readers who "sound out" new words in the course of processing a text. While we are generally not very concerned with "decoding" issues, this area cannot be ignored -- even for mature readers. For example; for native speakers of English,
Russian names with unusual spelling patterns are clearly much more difficult than English names—even when both are entirely novel. Doubtless, orthographic complexity plays a substantially greater role during the transition period when children are steadily adding words to their reading vocabulary and mastering spelling-to-sound correspondence rules.

Within the literature on "readability" there have been many studies which have shown that various indicators of orthographic difficulty and lexical familiarity, such as numbers of letters per word, number of syllables, and frequency, have a strong correlation with text difficulty (Bormuth, 1966; Klare, 1974-1975). In fact, in many of these studies these indicators were found to provide the best single predictor of overall text difficulty. However, these indices of orthographic complexity are only crude engineering approaches to the issues of understanding what makes text difficult. Although no one has yet devised any superior alternative measures of orthographic complexity, we suspect that such variables as word frequency is an index of the probability that the word in question is in the reading vocabulary of the child, whereas number of letters and number of syllables may well be crude indices of orthographic complexity. In any case, we feel that orthographic complexity is an important variable and we would like to determine how much of the difficulty associated with texts encountered during the transition stage can be accounted for by irregularity of the spelling-to-sound correspondence rules. Presumably as the reader becomes increasingly skilled, less and less of the difficulty will be determined by these factors and more and more by the "higher-order" aspects of textual material.
A fuller discussion of a number of these issues appears in Schallert, Kleiman, and Rubin (1977). Suffice it to say that we recognize that all levels of complexity contribute to the overall difficulty of textual material and in the formulation of our research plans we must keep in mind the orthographic level as well as the other more purely "textual" levels.

Lexical Level Variables

There is no clear notion what makes one lexical item more difficult than another. Historically, word frequency has been used as a measure of vocabulary difficulty. Thus, for example, Thorndike (1921a) discussed the issue and developed an early word count, The Teachers Word Book (1921b) to help educators do a better job of matching word difficulty to the knowledge of the student. But why should a "word count" measure word difficulty? The assumption behind frequency as a variable is that the higher the frequency of the word the greater is the probability that it is in the reader's reading and/or listening vocabulary. There are a number of difficulties with this assumption. One involves the fact that word counts are counts of surface lexical items and not meanings. Thus, for example, no distinction is made between tree in its usual meaning and when it is used as a technical term in transformational grammar—even though there is presumably considerable difference in difficulty. Another problem involves the normative characteristics of such word counts. People in different cultural milieux and with different interests have different linguistic experiences and thus different vocabularies. Thus, what may be frequent for one may be rare for another. To what degree do such differences account for differences in text difficulty for different children?
It is important to determine how much of reading difficulty can be attributed
to such variables. No child, however good his decoding skills or his general
comprehension skills, is going to be able to comprehend a text filled with
an unfamiliar vocabulary.

There are other problems with word frequency as a measure of
word difficulty. The most basic is that difficulty depends not only on
the number of times a child has encountered a word, but on its conceptual
difficulty. Thus, for example, most children will find the word *troglobyte*
unfamiliar and difficult. When told it means someone who live in a
cave, they should have little trouble learning the new word, since the
underlying conceptual form is readily available. On the other hand, the word
*left* in its political usage is quite different. Most children have encountered
this surface form relatively frequently, but will have considerable trouble
learning the new meaning, since the underlying concept is quite complex.

While there has been little work on this specific problem, the recent work
of Rosch on basic categories does begin to make contact with the issues (Rosch,
Mervis, Gray, Johnson & Boyes-Braem, 1976), as does the classic work of

One other problem associated with the traditional approaches to voca-
bulary difficulty is that it treats words in isolation, while the meaning
that get assigned to words is clearly dependent on the discourse context
in which the word occurs. Thus, for example, when we read of putting apples
in containers we are more likely to understand container to mean "basket" whereas
putting cola in containers is more likely to raise the image of bottles.
(c.f. Anderson & Ortony, 1975).
Thus, even at the level of vocabulary much basic research must be carried out before we can gain a true scientific understanding of "word difficulty." We need to know how to represent the semantic and/or conceptual structure of words. We need to know how meanings are constructed. It should be noted that to resolve the problems of the comprehension of text by young children it will almost certainly be necessary to have an understanding of these issues—a mere engineering approach will not work. For example, if one were to select vocabulary on the basis of the readability research, one might write text using words like adz or discuss the concept of time in modern physics since the word time is of high frequency. Clearly a deeper understanding of the underlying issues is required before it will be possible to give well motivated advice about the difficulty of texts.

Thus, we propose to evaluate the various models of lexical meaning that have been proposed (c.f. Schank, 1975, Norman & Rumelhart, 1975, Miller & Johnson-Laird, 1976) and attempt to relate them to textual difficulty and to the more general discourse variables to be discussed below.

Sentence Level Variables

There has already been an enormous amount of research on the effects of syntactic structures on comprehension; however, much remains to be done in this area.

Researchers in the readability area have shown that many surface syntactic features are related to text difficulty. The most frequently studied of these surface measures has been sentence length, but various other similar measures have also been investigated—number of clauses, number of prepositional phrases, ratio of conjunctions to pronouns (Bormuth, 1966; Klare, 1974 -1975). These variables tend to give correlations with text
difficulty in the .5 to .7 range; however, it is obvious that using these measures are not derived from a comprehensive theoretical analysis. It seems quite unlikely that replacing a sentence such as John likes chocolate ice cream and Bill likes vanilla ice cream with a shorter version such as John likes chocolate ice cream and Bill vanilla will make for easier comprehension. Once again, what is needed is an understanding of how and why particular aspects of the sentence structure make for text difficulty.

Another research tradition in the study of syntax has involved the study of the frequency of syntactic structures. The general approach has been to count the frequency of the patterns of children's spoken or written sentence, build texts based on these counts, and then test the comprehension of the texts (Bormuth, 1968; Jongsma, 1974; Peltz, 1973-74; Ruddell, 1965; Smith, 1971; Tatham, 1970). Most of these studies have found the more frequent syntactic structures to be easier to comprehend (Peltz, 1973-74; Ruddell, 1965; Smith 1974; Tatham, 1970). There have, however, been negative results as well (c.f. Jongsma, 1974).

There are a number of serious difficulties with this "syntactic frequency" approach. First, the effect is not even clearly established. The studies that gave negative findings used multiple choice tests, while most of those that gave positive findings used the cloze technique for measuring difficulty. Thus, the findings may not indicate that more frequent constructions are easier to comprehend, but rather that when filling in a blank children are more likely to use their own forms of expression than they are to use the forms of the original text authors. Even if this problem is ignored, the gathering of adequate data on frequency of syntactic
forms is a formidable undertaking. Most of the studies in this area have used the data of Strickland (1962), which is based on very crude surface analysis of syntactic patterns and a very limited sample. Solving these difficulties would require a subtle and practical procedure for analyzing the syntax of spoken utterances, something that is not easy to come by in the current stage in the development of linguistics. In addition, the fact that language is creative in the Chomskian sense means that even getting crude norms will require massive samples of language.

Even though there are many difficulties with the syntactic frequency approach to comprehension difficulty, it does have some mild virtue as an applied solution. It seems likely, that after the appropriate research is carried out, that the linguistic forms used by children of a given age group will be easier than those used infrequently. However, once again the relationship is likely to only be approximate, and so what is actually needed is serious scientific knowledge about which linguistic forms are more difficult and why.

There has already been considerable work in experimental psychology on the difficulty of various syntactic structures. Much of the early work consisted of experiments designed to show that Chomsky's linguistic theories could be used as a psychological model (Miller, 1965).

The early work suffered from a variety of conceptual and methodological difficulties (cf. Greene, 1972 for a review) and is of little interest here. However, out of this tradition have come a number of studies that do attempt to give a detailed account of the difficulty of a limited set of syntactic forms.
There have been a number of demonstration experiments which show that some linguistic forms are more difficulty than others (Coleman, 1964; Kintsch & Monk, 1972; Wang, 1970).

Unfortunately, these studies do not contribute to our detailed knowledge of the causes of syntactic difficulty. They all show that intuitively more "complex" syntactic forms are more difficult to comprehend than are intuitively "simpler" forms.

There is a number of experiments dealing with difficulties of particular syntactic constructions. Thus, for example, as a rule, right-branching and center embedded sentences are rated harder to understand than left-branching sentences (Schwartz, Sparkman & Deese, 1970); center embedded sentences are easier when relative pronouns are employed (Hakes & Cairns, 1970); that complement constructions are more difficult with the "that" deleted (Hakes, 1972) sentences with negations are more difficult than those without (c.f. Carpenter & Just, 1975, Sherman, 1975) etc.

All of these findings and many others are useful in as much as they give us a definite body of data on difficulties associated with particular forms. The problem, however, is that these results have been obtained in a rather piecemeal manner. What is required is a general theory of what kinds of syntactic structures are more difficult and why. Bever (1970) has provided a perceptual processing hypothesis which accounts for many of these results. Yet even Bever's principles do not always lead to a coherent account of all of them. Perhaps the most promising account of the effects of syntactic complexity comes from Kaplan's (1972, 1974) attempts to apply the augmented transition network (ATN) as a serious model of sentencing processing. (see also Wanner & Marotsos, 1974; Stevens & Rumelhart, 1975). It is our opinion that the development of a general processing
model such as this should be explored further since it offers the most promise for devising a precise measurement of the syntactic complexity of textual material. Such a model could be automatically applied to a text by a computer so as to generate both a global measure of complexity and a local measure of the complexity of each sentence and each phrase of a sentence (c.f. Kaplan, 1974 for his discussion of transient processing load).

Our discussion so far has focused on the syntactic structure of individual sentences. Of course, sentences do not occur in a vacuum. Just as word meaning often cannot be known without reference to the sentence in which the word occurs, so too sentence meaning often cannot be determined without reference to the general discourse in which it is embedded. In context, some of the syntactic effects mentioned above can be reversed. Thus, for example, although as a rule negatives are more difficult than positives Wason (1965) has shown that if the situation leads the hearer to expect a negative, negatives are easier than positives. Similarly, Olson & Filby (1972) showed that when passives are expected passives are easier than actives, although the reverse is normally true.

A number of studies have shown that semantic factors in sentences can drastically modify the effects of syntactic form. For example, Slobin (1966) showed that children found nonreversible passives much easier to understand than reversible passives, e.g., The flowers are being watered by the girl was easier than The cat is being chased by the dog. Stolz (1967) showed in a similar fashion that center-embedded sentences with semantic constraints were much easier than center-embedded sentence without semantic constraint.
Overall, it is clear that this type of experimental research must continue if we are to understand how syntactic complexity affects comprehension. However, to be of maximum benefit to the overall issue of text comprehension there needs to be some shifts of emphasis. We will need to study a wider range of syntactic forms. Better measures of comprehension must be developed. Both children and adults must be studied. More general models must be developed. The function of syntactic forms must be considered (why is a passive being used--for focus, for stylistic reasons, etc). The function of syntactic forms in discourse must be studied. This is clearly a fertile area of research and one of much benefit to the eventual issues of text difficulty in the transition grades.

Another quite different set of issues also arises at the sentence level due to the recent understanding that it is necessary to view sentences as much more than abstract syntactic structures. Austin (1962) caused a revolution in the analysis of language by pointing out that in order to understand actual linguistic utterances in communication situations it is necessary to understand what speech acts the speaker is carrying out. For example the speaker might be asking a question, giving an order, or making a statement. Austin proposed that utterances have an illocutionary force which conveys the intended speech act.

One of the difficult problems that speech acts pose for a model of language comprehension is that there is typically a very wide variety of constructions that express the same basic intention. For example one can ask someone to close the door by saying:

(1) Shut the door.

(2) Can you shut the door?
Clark and Lucy (1975) have proposed an information processing account of the understanding of conversationally implied requests. They suggested that the hearer first obtains the literal interpretation of the utterance. Then the hearer checks to see if the interpretation is appropriate to the context. If the literal meaning is appropriate then it is taken as the intended meaning. If the literal interpretation is inappropriate the intended meaning is inferred through a series of "rules of conversation" (Grice, 1975).

The analysis of language in terms of speech acts raises a number of interesting questions which must be addressed in the context of textual analyses. Since a wide variety of linguistic forms can be used by a speaker/writer to express the same speech act, are some forms more difficult to comprehend than others? To some degree the use of indirect speech and certain other indirect forms of expressing speech acts is a culturally relative factor. Do some children have special difficulty in understanding the implications of sentences like "Would you mind closing the door?" and if so, does this lead to any special difficulties in understanding the sorts of textual materials these children are expected to understand? Does the reader understand the intent of the author -- to amuse, to persuade etc. -- ? All of these issues are new and have not been studied in previous readability approaches to text analysis. Do these factors, in fact, play an important role in determining text difficulty?

A final series of problems at the sentence level are the various presuppositions and implications that must be understood before sentences are completely comprehended. Just and Clerk (1973) have studied the difficulty
of the presuppositions and implications that must be understood before sentences are completely comprehended. They assume that for a sentence such as John managed to find his hat there is presupposed John tried to find his hat, and the implication is that John found his hat. Harris (1974) studied a variety of verb classes in complex-complement sentences and showed that adult subjects made the expected inferences. For example, in a rating task they indicated that the sentence John regretted that Bill was absent implies that the speaker believes Bill was absent, while the sentence Harry pretended that Bill was absent implies that Bill was not absent. In a second study Harris (1975) studied children's acquisition of this type of understanding and found that some of the classes of verbs were not understood until quite late. Once again it is clear that the problems relating to the presuppositions and implications of sentences is a complex and rich topic that must be studied before we can give an adequate account of text.

Discourse Level Variables

In our attempt to analyze textual material, we, of course, cannot stop at the level of sentences. We have already pointed out that words must be understood in sentences in which they occur and that sentences must be understood in the discourse context in which they occur. Most of the work by psychologists and linguists over the last 20 years has focused on words and sentences; serious work on discourse factors in comprehension and text difficulty is in its infancy. We believe, however, that not only are discourse variables essential to our understanding of comprehension of textual material, but that indeed these are the critical variables in the analysis of textual material. The work on discourse variables can
conveniently be divided in two parts: the primary experimental work coming from psychology and the more formal work on discourse variables deriving from linguistics. We begin with a discussion of the psychological approaches.

**Psychological approaches.** As we have already said, traditional research in the area of readability has used variables such as word length and frequency as the basic elements (Klare, 1974-1975). Consequently, it has failed to reveal any contribution that discourse level variables might make to text readability. That discourse level variables cannot be ignored is probably obvious once stated.

Within experimental psychology there have only been a few studies directly relevant to the comprehension of discourse level phenomena and so it will be necessary to include in this section a number of studies that used memory tasks which are only an indirect measure of comprehension.

Deictic aspects of language can play a considerable role in comprehension. Brewer and Harris (1974) and Harris and Brewer (1973) have shown that if subjects are asked to remember sentences isolated from the actual time, place, person, and discourse context the deictic words are much more difficult to remember than the nondeictic words. In sentences with appropriate contexts the deictic words were little different from the nondeictic words. Thus, use of the deictic words away from their normal context seems to render them uninterpretable and hard to remember.

Much basic work on the comprehension of deictic elements remains to be done. For example, how does a hearer use knowledge about the nonlinguistic aspects of the world to interpret deictic words in sentences? This may be an important problem in reading comprehension since the
problem of deixis is somewhat more complex in written material than in spoken discourse. In written discourse the knowledge about time, place, person, and context must be explicitly presented or implied in the prose itself since the individual attempting to comprehend the material is not actually present and aware of the time, place, person, and context. Written discourse also makes frequent use of deixis to refer to aspects of the prose itself. Thus, the writer refers to this argument or to the latter point. Everyone who has edited written material knows that the use of the vaguely specified this can have a strong effect on the comprehension of prose. These matters must be investigated with transition level readers and further discussion of them can be found in the report on Oral and Written Language.

One of the crucial devices that holds text together is anaphoric references--aspects of one sentence referring back to someone or something in an earlier sentence. Lesgold (1972) showed in a probed recall task that sentences with pronominal co-reference (The aunt ate the pie and she was senile) frequently functioned as a unit while sentences without co-reference (The aunt ate the pie and Alice was senile) acted more like two separate units.

devilliers (1974) found a similar effect with determiners. He found that a text with definite articles in it (showing co-reference) was better recalled than the same text containing indefinite articles. Bormuth, Manning, Carr, and Pearson (1970) tested children's knowledge of anaphoric reference by a question asking technique. For example, the subject would hear a pair of sentences such as The boy fell off the steed. He fractured his arm, and then would be asked Who fractured his arm? Bormuth et al. tested a number of different forms and found a substantial range of
difficulty for children. Haviland and Clark (1974) looked at comprehension in terms of given and new information. They found that the second sentence of the pair Andrew was especially fond of beer. The beer was warm. was slower than the same sentence in the pair We got some beer out of the trunk. The beer was warm. Again the state of the area is such that we need more conceptual and empirical work before we can begin to understand the full complexities of the various anaphoric devices in text comprehension.

Perhaps the most powerful discourse variable which has received attention is the role of importance in comprehension. Clearly, importance is discourse level phenomenon. What is important in one context is unimportant in another. It has long been known that more important information is better recalled than less important information (e.g., Bartlett, 1932; Binet & Henri, 1894; Gomulicki, 1956; Johnson, 1970). This recall advantage indicates that readers somehow extract, and given special prominence to those more important aspects of the text. Thus, if beginning readers are to become mature readers, they must learn those characteristics of texts which reflect importance. We have proposed a series of studies on this general topic in the section of task analysis.

One of the functions of a theory of content structure is to provide a theory of importance. Presumably, those aspects which are most important play salient roles in the structure of the content. Thus, accounts of discourse structure such as that by Rumelhart (1975, 1976), Schank & Abelson (1975) etc. have focused on summarization as a key dependent variable in their research. The issues of content structure play such a large role in our proposed research that we have devoted an entire section to it. Thus, we put off further discussion to that section.
Although the linguistic work on such discourse variables is limited, still it is more substantial than the psychological literature. It will be useful, at this point, to discuss this work before proceeding with a discussion of effects of the global organization of textual content and other, more extra-linguistic variables.

**Linguistic approach.** An important and intriguing question both for research on comprehension and for the analysis of texts is to what extent, and in what ways, are purely linguistic devices used to provide clues to the content structure of the text? Though the study of this question has only recently begun (at least in this country), there are clear indications from linguistic research that certain linguistic phenomena function more or less directly as signals of various aspects of text or discourse structure.

There are four trends in linguistic research relevant to this question. The oldest of these is the theory of "functional sentence perspective" (FSP), most closely associated with the Linguistic Circle of Prague, but including similar work by Halliday (though the term FSP is not often applied to his work), and most recently work in the United States by Susumu Kuno that combines his view of FSP with certain aspects of transformational grammar. For a representative sampling of FSP, see the various issues of *Travaux Linguistique de Prague*, the papers in Danes (1974), especially those by Firbas and Daneš; and Sgall (1975). Typical related work by Halliday appears in Halliday (1970, 1974) and Hasan and Halliday (1976). Kuno's work can be seen in Kuno (1972, 1975, 1976).
Research in FSP goes back to the work of Mathesius and other linguists of the Prague School on the syntax of Slavic languages. Slavic languages generally have more highly developed case-marking systems and, probably as a result, considerably more freedom of word order than English. But in studying variations in word order, Prague linguists noticed that word order is constrained by discourse factors. Elements expressing "known" or "old" information generally precede elements expressing "unknown" or "new" information. The primary difference in this respect between English and Slavic languages is that in English word order changes are brought about by rules that also change grammatical relations; passivization, for instance. Firbas (1964, 1966, 1971) has written in detail on the relation of English word order to information distribution.

From these origins there has arisen an extensive literature on discourse notions like "topic", "focus", "comment", "given/new", "new/old information", and "communicative dynamism". Although there are differences among the various proponents of FSP, there is a common theme: close attention to language as a communication system (as opposed to the recent American bias toward description of language as an abstract formal system); in particular, there is a growing interest in reflections in linguistic forms of discourse properties. Generally speaking, FSP linguists have a three-level approach: the level of semantic structure, the level of grammatical (roughly, syntactic) structure, and the level of the "organization of the utterance" (see Danes (1974) for discussion). It is the third level that is described in terms of discourse-functional properties and their reflection in grammatical form, and it is this area that relates directly to the matter of the structure of discourse, hence of texts. The interaction of such matters with questions of the structure of content needs close investigation.
We expect that many of the concepts of FSP theory, especially "topic", "new/old information" and the insight that these matters at least partially condition some syntactic properties, will be of great importance in understanding the role of linguistic form in the interpretation of content structure. As we have seen, psychologists have already begun an investigation of the "given/new" distinction and its role in comprehension (Haviland & Clark, 1974).

The main problem with this approach (and all theories that deal with these matters) is that crucial, intuitively appealing notions like "topic", "given/new" and so forth, even though central to the story, are not sufficiently clear and explicitly defined or described, so that it is not clear how to apply them in any but the simplest cases. This makes it difficult to test the predictions. While we find Kuno's (1972) innovations intuitively very appealing, there are serious problems in evaluating them. The crucial notions "theme", "contrast", "exhaustive listing" and "neutral description" are not adequately defined, and are consequently difficult to test. The immediate goal for relevant linguistic research in this area should be to develop more precise characterizations of central concepts, and some sort of replicable diagnostics for their occurrence. Some work along these lines has begun (see for example Chafe (1976), Keenan & Schieffelin (1976), and Li & Thompson (1976) for exploratory work). At the same time, the FSP framework should be compared as a model of discourse with other theories of text structure, such as those of Text Grammar, Grimes, Rumelhart, and so on.

A second theory, more recent than FSP, is the theory of Text Grammar (TXT). This theory is concerned directly with aspects of content structure.
Based on the observation that many linguistic phenomena cannot be
dealt with insightfully at the sentence level, proponents of TXT take
the position that it is a mistake to take the sentence as the basic unit
and domain of linguistic theory; rather, the text or discourse should be
basic. According to van Dijk (1972), discourse is the only justifiable
natural domain of an empirically adequate theory of language. The domain
of sentence-grammar is to be entirely subsumed in TXT, and no subpart of TXT
will be isolatable in any natural way as a sentence grammar. Though TXT
is still in a relatively embryonic state, more a program for research
than a full-fledged linguistic theory, it is possible to discern in it
two main variants: a formally-oriented variant, and a functionally-oriented
variant.

The former is a tendency to carry over from sentence to discourse domain
the view of linguistic theory as a mechanism for generation of well-formed
formal objects, independent of speaker, hearer, and communication. The goal
of early TXT is to characterize "coherent text" generatively, by providing
formal rules to generate the set of well-formed texts. The second variant
of TXT is more functionally oriented. This work is even more programmatic
than formally-oriented TXT. Its main proponents are Klein (1972), Brinker
variant is also concerned with the structure of discourse, it is orientated
less toward formal apparatus for the generation of text and more toward the
construction of a model of communicative competence. Here linguistic units
are not seen as isolated formal objects, but with respect to the socio-
communicative context in which they are used by speakers. Hence the starting
point for functional theory and analysis of text grammar is linguistic
communication in interactive situations. But even this approach has a role for the analysis of discourse structure. Sandif (1973), for example, argues for a view of text as either sequences or hierarchies of speech acts. Recent work by van Dijk (1975, 1976) seems to give a larger role to functional considerations, though he is clearly still concerned with formal characterization of well-formed texts.

The promise of this work for the analysis of the content of discourse will be discussed in the following section. It may provide useful descriptions of the relation between content structure and linguistic form, though work in this area is not yet very specific.

In this country, linguistic work relevant to text analysis is connected with two separate linguistic theories. The first is in work on discourse structure in the familiar framework of tagmemic theory, associated primarily with Kenneth Pike and Robert Longacre, and their co-workers. (see Pike (1964), Longacre (1968, 1972)). In this framework the attempt is made to apply the slot-filler-tagmeme mode of syntactic description to the description of discourse, at the level of sentence, paragraph, and discourse, yielding a hierarchical analysis in terms of form and function. Attempts are made at classifying types of discourse ("narrative", "hortatory", "expository", etc.; see Longacre (1972), Chapter five) as a first step toward analyzing the form/function relation in discourse, and within this taxonomy to relate discourse function to grammatical form. This approach includes an abstract "deep structure" of the discourse, and rules relating this structure to grammatical structure in terms of intersentential relations and grammatical devices, making use also of case-like concepts like "agent", "instrument", etc. The bulk of tagmemic work on discourse structure has
been on languages of New Guinea and the Philippines, but the insights embodied in their analyses may be useful for the analysis of English texts. The examination of English discourse in contrast with other languages may point up discourse devices unique to English.

Work on the relation between discourse structure and linguistic form—syntactic properties in particular—has been increased recently in generative grammar. It has become clear in this framework, as well as the ones discussed earlier, that some syntactic properties cannot be fully described without reference to discourse and content factors. This includes not only processes of pronominalization and ellipsis (Ross, 1969; Grinder & Postal, 1971; Morgan, 1973; Sag & Hankamer, 1976), but also discourse conditions on syntactic rules (Hooper & Thompson, 1973; Green, 1974, 1976; Morgan, 1975; Cary, 1974; Ziv, 1976; and others). This work indicates that syntactic properties can serve as clues to discourse structure at various levels.

The most ambitious generative approach to discourse and syntax that grows out of this research is contained in unpublished work by G. Lakoff on the theory of "dual hierarchy grammar". In this theory syntactic relations are treated as predictable from a combination of discourse and semantic factors. This theory is yet in its embryonic stages.

Work being done in the generative framework is by far the most explicit and detailed being done on matters of linguistic form. What it lacks is a coherent theory of discourse on which to base its descriptions.

One of the first requirements for our theoretical research is to compare these approaches and to answer such questions as: on what points do they differ substantively? Do they yield useful analytic procedures?
How can they be improved? An attempt must be made to bring together the detailed description of form of the generative work and the models of discourse of the others. For the time being, until a unified approach can be constructed, the most promising strategy is to continue detailed analysis using whatever rough-and-ready discourse description is most convenient; then once content analysis procedure (and content-analyzed texts) are available, work can begin on detailed study of the relations between content and linguistic form.

In the meantime, the most pressing need is for more explicit theories of discourse. There are many anecdotally illustrated theories of discourse available, but in them there is widespread vagueness, inconsistency, and confusion on the exact nature of their central notions. (See Chafe (1976) for a brief but revealing discussion of the generally confused state of affairs.) Even such an intuitively simple notion as "contrast" has never been defined or modelled, and is likely to be surprisingly difficult. We do not mean to say that none of the existing approaches are correct; only that it is very difficult to subject them to empirical test, and to "raw out" their implications in detail, until they are made more explicit and rigorous. At this point what is needed is not so much grand theorizing as detailed theoretical and empirical analysis, for the sake of increasing the level of precision and rigor involved. This entails both an increase in conceptual precision and the development of explicit formalism to reflect that precision. The necessary research lies on the fading border between linguistics, cognitive psychology, and artificial intelligence, i.e., in "cognitive science". Realistic models of discourse will necessarily draw on all three, in that problems in the representation of knowledge, models of comprehension, and linguistic semantics and pragmatics will all
have to be dealt with in the construction of discourse models of the necessary precision.

The first step in this direction is to devise a program of research that inherently demands a high level of explicitness and precision, so that investigators will know immediately when the model lacks clarity. One good strategy, besides the usual sort of theoretical research strategy, is the use of computer modelling techniques as a research tool. Beginning efforts will no doubt require much human intervention, but the constant goal should be the reduction of such human guidance. A good beginning would be to attempt a completely explicit formulation--computer modelling, if possible--of the simplest discourse notions, like "contrast" and "exhaustive listing". Given some degree of success with these, more difficult concepts like "topic" and "new information" can be attacked, including attention to detailed questions like: Can there be more than one topic per sentence? If so, under what circumstances? Is it possible to have several topics simultaneously at different levels of discourse? If so, are they distinguished in any way in the linguistic system? How does the notion "topic" relate to syntactic phenomena of "topicalization"? What is the nature of the interaction, if any, between "topic" and "old information"? Are the two distinct? We fully expect that the number of questions like these that arise will greatly exceed the answers arrived at; such a result will be both a sign of progress and an indication of the present state of ignorance. In this research those aspects of linguistic form should be a central factor that have been shown to be related to discourse structure, and new ones should be looked for, as they presumably will provide one of the best (and most easily controllable) tests of the appropriateness of the theoretical concepts being investigated.
Content Level Variables

In addition to the various ways that text can be worded, structured or organized, it is also possible to study the structure of the content of the discourse itself. A body of work focusing on the logical and narrative structure of textual material has been accumulating over the past several years. In this section we sketch the major boundaries of the work and indicate the major directions our own efforts in this regard will take.

For the purpose of the analysis of text content, it is necessary to develop and test effective descriptive techniques for content analysis. There is a need for research on two fronts: first, further research on the role of content structure and inference in text comprehension with a view to developing a model of comprehension; second, research on practical procedures for the descriptive analysis of texts.

Test structure and comprehension. The simplest and oldest notion of text structure is that the theme is the unifying element that runs through the next and binds the elements together. When the theme is vague and difficult to discover, as in studies by Dooling & Lacman (1971), Transford & Johnson (1972), and Dooling & Mullet (1973), subjects are unable to benefit from the theme relatedness of the passage, unless they are cued onto the theme with a title or picture. Without such cuing, subjects find texts impossible to comprehend and difficult to remember. The manipulation of the theme by choosing the title given to an ambiguous passage can dramatically alter the interpretation (Schallert, 1976).

Theme relatedness is a powerful determinant of recall. Subjects tend to remember the main theme of a story and to forget details or peripherally related events. Recall may therefore be characterized as an abstractive process. For example, Comulicki (1956) investigated the immediate recall
of prose passages from 1 to 200 words. He found that although subjects were able to recall the shorter passages verbatim, they were only able to recall the more important aspects of the longer passages. Further, he found that omissions were by far the most common error type, and concluded that memory for passages was most accurately described as abstractive process. In fact, when judges were given both recall protocols and deliberately written abstracts of the same passages, they were little better than chance at distinguishing between them. Theme relatedness determines the importance and memorability of information in discourse. Johnson (1970) measured importance of pausal units in discourse simply and directly by having subjects rate their importance. Rated importance was found to be a powerful determinant of recall.

Importance has since been systematized by the development of text structures from which measures of importance could be derived. Meyer & McConkie (1973) used quite a simple method of discourse structure analysis. They had graduate students outline a passage, and then converted the outlines to tree structures. From these tree structures, three measures of the importance of an idea unit in the structure of the passage were developed: a hierarchy depth score, which measured how high in the hierarchy the unit occurred; a units beneath score, which measured the number of units which were beneath the given unit in the hierarchy; and a combined hierarchy score, which combined the two above measures, equally weighted, into a single, unified measure. Significant effects upon recall were found for all three measures. Further, when significant effects of serial position and rated importance were found, these turned out to be
largely due to the correlation of those factors with hierarchical importance. They also found that if a unit was recalled, then there was nearly a 70% chance that the unit which occurred immediately above it in the tree was also recalled, although, overall, recall was only about 23%. Further, combined hierarchy score was positively correlated with stability of recall across two recall trials. The effect of importance has been replicated using more formal text structures (e.g., Kintsch, 1972, 1974; Meyer, 1975, 1977.) Kintsch has replicated the results of Meyer & McConkie (1973), using his more formal and objective propositional description. Kintsch's (Kintsch & Keenan, 1973) propositional rank is essentially equivalent to Meyer and McConkies' hierarchy depth score, and Kintsch's counting of descendant propositions is analogous to Meyer and McConkies' units beneath score. Essentially similar results have been reported using Rumelhart's (1975) story grammar as the structural representation (c.f. Mandler & Johnson, 1977; Rumelhart, 1976; Thorndyke, 1975).

The body of research on this topic is growing. We expect that further work along these lines will follow naturally from our development of adequate procedures for encoding textual materials. We turn now to a discussion of our proposed work on this topic.

**Descriptive analysis of text content.** There are many content variables that must be evaluated and coded in the descriptive analysis of texts. Among these are the author's intentions at various levels, inferences that must be made in comprehending the text, and the structure of the content of the text.

For practical descriptive purposes, rough intuitive typologies will probably suffice for description of global aspects of authors' intentions like
subject matter, purpose (description, explanation, persuasion, practice, entertainment, etc.), and genre (narrative, instructions, dialogue, etc.).

With respect to problems of inference it will be necessary to distinguish among several types and levels. Types include inferences that do and do not require substantial background knowledge outside the text, and inferences involved in the interpretation of non-literal language. Another possible type of inference that needs investigation is the matter of "textual postulates". The question here is, are there "postulates" peculiar to written texts distinct from the "conversational postulates" that play a role in conversational interactions (see Gordon & Lakoff (1971), Searle (1975), Morgan (1977) for discussion)? If so, are there different "textual postulates" for different kinds of texts?

Levels of inference range from inferences that must be made about the author's global intentions in the text, to proposition-level inferences involved in the interpretation of particular sentences, or in inferring connections between the proposition-level units of the text.

An important practical goal of research on these questions should be the development of replicable procedures and techniques for identifying and coding the various kinds of inferences involved in comprehending a text. Frederiksen (1975) has already made one such proposal. It would also be very useful to develop relative measures of the "explicitness" of text; that is, of the amount of inference that must be made in comprehending the text.

The third matter, the structure of the content of the text, is clearly a very important factor of text analysis. The basic idea of text structure is not a new one, but goes back at least to the work of the Russian Formalists, Propp (1928) in particular, on narrative structure. There have
even been occasional attempts to extend the apparatus of transformational
grammar to the analysis of narrative structure, but this work has received
very little attention in linguistics. Recently, however, the notion of
text structure has begun to receive attention in psychology and, concomitant
with the growth of interest in discourse in linguistics; see for example
Frederiksen (1972, 1976), Grimes (1975), Rumelhart (1975, 1976), Schank (1975),
and van Dijk (1975, 1976). This work varies in rigor, explicitness, generality,
and on many details of analysis. But it is likely that the common
threads running through this work will lead to useful descriptive tools.

The common threads are these: (1) a multi-level approach to the
organization of the content. This includes at least a propositional
level, consisting roughly of the sentences of the text and proposition-level
inferences, represented usually by a system reminiscent of a Fillmorean
case system (see Fillmore, 1968) or "schemata" (see Rumelhart & Ortony, 1976,
for discussion), (2) a higher level of organization of the text, usually
hierarchical, except perhaps in the case of Schank (1975), where the
structure consists of "causal links" among the proposition-level units
(explicit and inferred) of the text. The details of this higher level
differ; Rumelhart (1975), for example, analyzes story structure in terms
of two paired sets of rules: a set of syntactic rules specifying the
structure of the story in terms of functional categories like "setting",
"episode", "event", and so on, and paired with each syntactic rule a
"semantic interpretation rule" specifying the relations ("CAUSE", "ALLOW",
"MOTIVATE", etc.) among the categories specified in the corresponding
syntactic rule. Grimes (1975), on the other hand, treats these matters
separately, as "non-events" and "rhetorical predicates". Some researchers
(e.g., Meyer, 1975) supply explicit analytic procedures; in others the
procedure is implicit. Some (e.g., Rumelhart (1975, 1976), Schank (1975), van Dijk (1975), offer attempts at explicit summarization rules. It is generally at this level of content structure, in terms of relations between elements of the content and/or height in the hierarchical structure, that matters like importance and cohesion are treated. (3) The role of inference is crucial to this approach, both in "filling in the blanks" by inference at the propositional level, and in inferring relations among the higher-level units of the text.

It should be possible to develop from this work at least informal procedures for analyzing and describing the content structure of texts. The ideal theory of content structure, of course, would be one which provided a procedure for analysis which did not rely heavily on the analyst; rather, the theory would be an algorithm that would take as input a text, a context, and a model of knowledge of the world, giving as output a representation of the organizational and logical structure of the text. Given the immense linguistic and psychological problems that must be solved -- a complete theory of linguistic semantics, and a coherent way of representing world knowledge -- we do not expect this goal to be reached in the near future. In the meantime, existing research on text structure should be mined for analytic techniques by, first, attempting to extract from each approach a procedure for analysis of text content; second, by beginning to apply each resulting procedure to selected reading texts, from references obtained from text sampling research; and finally, by comparing results of various approaches and modifying procedures where necessary. Such modifications should be motivated by at least four considerations: generality, replicability, level of detail, and relative merit. (a) Generality: does
the procedure yield useful and perspicuous analyses of a wide variety of texts? Rumelhart (1975), for example, seems appropriate only for stories; others (e.g., Grimes, 1975; Frederiksen, 1975) are more general. But it remains to be seen whether there is a "universal grammar" for texts. The most general scheme may fail to give insight into important differences between texts. It is conceivable that what is more useful is a repertoire of analytic schemata for various types of texts as, for example, proposed by Rumelhart (1976). (b) Replicability: does the procedure produce essentially similar results when used by different analysts? There is some evidence of replicability in the literature (e.g., Meyer, 1975), but the matter deserves more study. (c) Level of detail: does the theory yield procedures with a good level of detail in analysis, or only very general aspects of structure? (d) Relative merits of the resulting analyses: are differences substantive, or merely terminological? Are there simple translation algorithms from one procedure to another, showing that they are equivalent?

This work should not be done in vacuo, of course, but should be informed by theoretical research on comprehension and the representation of knowledge. The result of the research outlined here should yield: (a) best analytic technique (or repertoire of techniques) for analysis of text content, (b) paper discussing the relative theoretical and practical merits of the various theories and procedures, (c) an initial set of content-analyzed texts, for use in research on the relation between content structure and linguistic structure of the text, (d) work in this area, together with research on the relation between text structure and linguistic structure,
should lead eventually to procedures and measures for evaluating the "transparency" of text structure; that is, the degree to which the intended structure of the text is easy or difficult for the reader to discern, and for evaluating the relative complexity of text structure.

Other Extra-linguistic Variables

As we have emphasized throughout, aspects of the text which may appear as important determiners of difficulty on one level of analysis can always be over-ridden by variables at a higher level of analysis. In the last analysis, the difficulty of a piece of textual material does not lie solely or perhaps not even primarily in the text. Rather, it is an interaction between the textual material, and the skills, knowledge, motivation, and expectations of the reader. On this view, understanding is really the construction of an interpretation of a text consistent with the clues provided by an author. Those clues may be more or less difficult for any given reader to interpret. This difficulty can depend on numerous variables in addition to the actual sequence of symbols normally taken to constitute the text. In this section, we turn our attention to a few of these variables.

The role of illustrations. No one can doubt that illustrations play an important role in comprehension. In early reading materials illustrations probably carry the bulk of the content of the passage being read. However, illustrations are far from a semantic crutch for the illiterate. They often play a critical role in the comprehension of textual material. There is little information available pinpointing the precise role such information plays in comprehension. There are, however, several studies which demonstrate that it can play an important role. For example, Matz & Rohwer (1971) reported that the addition of redundant pictures increased the comprehensibility of a passage presented aurally to fourth graders. The increase was especially
significant for low-SES black students. In another study, Arnold & Brooks (1976) have demonstrated large effects on recall and ability to generate correct inferences when a text was accompanied by pictorial organizing--rather than verbal organizing--material. In another very nice demonstration Braisford & Johnson (1972) showed how pictorial information can convert a totally senseless passage into a very meaningful one. To date, most of these experiments have been mere demonstrations rather than contributions to a detailed account of the role of pictures in textual structure or in comprehension.

One way of remedying this lack of information would be to examine the role of illustrations in the comprehension of naturally occurring written material. A first step would involve locating texts where explanations are accompanied by pictures, charts, or graphs. One could then examine the shift in types of pictures found in texts used by elementary grade children (K thru 8). More specifically, the following questions would be addressed: (1) how much information is represented in the pictures (i.e., how well can one understand the author's intent simply by looking at the pictures); (2) how much information in the pictures (graphs, charts) is dependent upon the text for clarification; (3) how much information in the text is dependent upon pictures, graphs, or charts to become comprehensible; (4) what proportion of instances occur where the text is redundant and/or where the pictures are redundant? It seems reasonable to expect that younger children will be exposed to more cases in which text depends on pictures than cases in which pictures depend on the text. Also, the proportion of redundant text-picture situations should decrease with increasing grade level.
Following an analysis of what children naturally experience in terms of the relationship between linguistic and non-linguistic information, experimental manipulations can be made which vary written presentations along the four dimensions implied above and comprehension could then be measured. A comparison of particular interest would be between self-contained, non-illustrated text, text with entirely redundant illustrations, and illustrated text where neither text nor pictures carries the whole message (e.g., comics). It seems reasonable to expect that young and poor readers, who have not yet learned to extract the full meaning from "decontextualized text," will be adversely affected by conditions where the intended message is not carried fully by accompanying illustrations.

There is also evidence that illustrations play a role in providing clues as to the structure of the content as well. In a preliminary analysis in which Rumelhart's (1975) story grammar was applied to several children's stories it was found that the illustrators of those stories had elected to provide one picture for each EPISODE (in the sense of the grammar) of the story. This suggests that we should look closely at those places in the text illustrators choose to place their pictures and attempt to relate these places to significant units of the story as defined by the various models of textual structure discussed in the previous section.

The role of interest. Textual difficulty depends not only on the marks written on the page, but also on the attitudes a reader brings to the task of making sense of them. As a rule, oral language is used very purposefully. We speak to others and attempt to understand what they say not because we are instructed to do so, but because we intend our speech acts to accomplish certain results. In short, we speak and listen as communicative acts.
We read because we want to know about what is written. We write so as to express certain of our thoughts. For the child learning to read, reading is often not a communicative act. Rather, children often read without any particular concern for the message under discussion, but because they must read to fulfill certain school requirements. If, children, like adults, could approach reading as a communicative act—a means of finding out what s/he wants to know, attitudes toward the reading process might well change and the child may well be induced to make "the effort after meaning" required for comprehension to occur.

Perhaps the most obvious content variable which has been manipulated in this regard is the child's interest in the topic of the textual material. Presumably, if the author is trying to communicate information about which the reader has a strong interest, the reader will more readily seek that information. Thus, considerable effort has been devoted to the assessment of children's interests in the content of their texts.

What is surprising is that this line of inquiry has not led to much research on the way in which children's interests influence comprehension. Blom, Waite, and Zimet (1970), who have done the extensive content analysis of children's reading primers, have commented "What is needed is an investigation into how content actually affects children's attitudes and their acquisition of reading skill" (p. 219).

Much of the literature aimed at determining whether interesting material facilitates children's reading comprehension contains serious methodological problems. First, there has been a tendency to measure children's interest in a topic after they have read a page on that topic (Bernstein, 1955; Shnayer, 1967). This procedure compounds the reading comprehension measure with the interest assessment procedure since children may report more interest
in a passage that they have just understood. A better procedure would be to assess children's interest independently of any reading material and independently of the reading task. If superior performance is associated with high-interest material, then the rival interpretation that children prefer material that they do well on can be eliminated.

A second approach to the interest assessment problem has been to assign material to children based on normative data on children's interests (Dorsel, 1975; Klein, 1969; Stanchfield, 1967). While it is the case that children's interests are highly sex-typed (see, Asher, 1975; Tyler, 1964), research in the area also indicates a fair amount of variability of interest in a topic among males or females. The approach that should be adopted would be to individually assess children's interests and assign passages which are individually appropriate.

A third problem in the literature of children's interests and reading comprehension is the type of reading comprehension measure employed. Studies have used reading achievement tests specifically developed for each study with no prior demonstration of test reliability or validity. In many cases, item selection appears to have been arbitrary.

Finally, most of the research in this area has not sampled from a wide array of reading topics. Frequently, two passages are used, one of which is supposedly high-interest and the other supposedly low-interest (e.g., Klein, 1969).

A recent study by Asher & Marcell (1974) provided an improved methodology for studying the effect of reading content on children's reading comprehension. Children's interests were assessed independently of any reading material by having them rate photographs on an interest scale.
The photographs represented a wide array of topics. One week later, each child received from a second experimenter an individualized set of six passages. Three passages corresponded to the child’s three highest rated topics and three corresponded to the child’s three lowest rated topics. The cloze procedure (Taylor, 1953) was used as a measure of reading comprehension since the procedure provides a replicable and objective method for generating test items, and since cloze test results correlate highly with results from standardized achievement tests (Bormuth, 1967; Rankin & Culhane, 1969).

This methodology has been employed in several experiments to date. The first (Asher & Markell, 1974) was conducted to assess the extent to which sex differences in reading comprehension might be a function of the interest of the reading material. Elementary school boys often are found to read more poorly than girls on standardized tests (Asher & Gottman, 1973; Bates, 1961). In addition, boys are much more likely to be found in remedial reading classes (Blom, 1971).

Results of the initial experiment were striking. Boys’ reading comprehension was considerably higher on high- than low-interest material. Girls did slightly better on high- than low-interest material, but the difference for girls was not statistically significant. As for sex differences, the data indicated that girls scored significantly higher than boys on the low-interest passages, but on the high-interest material, the sex difference was eliminated. These data provide a powerful demonstration of the effect of the interest in material on reading comprehension. The oft-cited sex difference in reading comprehension appears to be at least partly a function of the nature of the material children are given to read.
The final experiment in this series (Asher, 1976) also found that girls as well as boys benefited from being given high-interest material. The major purpose of this experiment was to assess the effect of interest on black and white children's comprehension, but sex was also included in the analysis. The results, with regard to sex (see below for a discussion of race and interests) indicated that both boys and girls achieved higher cloze scores on the high- rather than low-interest passages.

What accounts for the discrepancy between the initial experiment and the subsequent ones? Why did girls do better on high-interest material only in the latter two experiments but not the first? One possibility is that children in the initial experiment were developmentally different from those in subsequent experiments. School-administered reading achievement data were available for each sample. All three experiments were conducted in the same school district so data from the same achievement test were available (Educational Development Series of the Scholastic Testing Service). An interesting pattern emerges. In the first experiment, girls outperformed boys on the standardized tests. In the subsequent experiments, no sex difference was found.

Fifth grade is a transitional period with respect to sex differences in reading comprehension. It appears that sex differences are rather consistently obtained with younger children, typically not found with older children, and inconsistently obtained in fifth grade (Gates, 1961; Hughes, 1953; Stroud & Lindquest, 1942). It may be that the effect of interest on comprehension interacts with the developmental level of children. In the earlier elementary years when girls are outperforming boys in reading they may be strongly oriented toward doing well in school. The content of reading material may
have less effect because they are seeking to do well regardless of the type of material they are asked to read. Later a shift may occur. Girls may cease seeking as strongly to achieve for its own sake or for external approval, and thus, become more responsive like the boys to the type of reading material they are asked to read.

What is suggested, then, is that when elementary school boys and girls perform equally, it may be because girls have become less motivated to excel, regardless of the nature of the task they are given. Thus, in the samples when boys and girls performed equally on the school-administered standardized test, both sexes were facilitated by interesting material. In the sample where girls achieved higher standardized test scores, boys but not girls were affected by high-interest material. This interpretation is highly speculative but testable. One approach would be to study the effect of interest developmentally. Children below fifth grade (e.g., third grade) should show stronger effects for boys than for girls. Older children (e.g., sixth grade) should show effects for both sexes. Another approach would be to sub-classify fifth grade children based on achievement scores or relevant motivational measures and determine whether certain groups of children show the effect of interest and others do not.

Finally, we turn to a discussion of some unanswered questions. It can be concluded from the available data that children comprehend more of high- than low-interest material. It has not been established, however, why this effect occurs. The preceding discussion about sex differences implies one type of explanation; namely, that children comprehend more of interesting material because they are more motivated to engage in the task. Aiken & Marwell (1974), in attempting to explain their initial
findings, invoked one type of motivational account. Citing data that boys and girls view reading as a feminine activity (Kagan, 1964; Stein & Smithells, 1969), Asher and Markell suggested that high-interest reading material, through its association with traditionally masculine topics (e.g., basketball, race cars), might serve to make reading a more sex-appropriate activity. This explanation has the virtue of explaining both the fact that boys were facilitated by high-interest material and that girls were not. Girls, after all, would not need high-interest material to define reading as sex-appropriate. The problem with the explanation is that it doesn't handle the finding from later experiments that girls, too, performed better on high-interest material. A more general motivational account is required.

There is also the possibility that the interest effect is not a motivational phenomenon at all. Children may try just as hard on low-interest as on high-interest material, but due to knowledge constraints do less well on the low-interest material. Two types of knowledge variables can be suggested. First, children may have less familiarity with the vocabulary of low-interest topics. Hence, when attempting to replace deleted words on the cloze task they may have greater difficulty when the material is less familiar. A second possibility is that children have more elaborate and differentiated cognitive structures with respect to topics they are interested in. Interest in a topic implies that the reader has schemata about that topic which help him organize material, infer the writer's intentions, anticipate future discussion in the text, etc. The greater availability of such schemata would certainly facilitate cloze performance on high-interest passages.
In the real world of reading, knowledge and motivation are confounded. Attempts to stimulate interests involve imparting new information as well as motivating a student to explore a new area. New information, or old information presented in a new light, does much to spark interests. Although motivational and knowledge factors are confounded in everyday life, it would be instructive for theoretical, and perhaps for practical reasons as well, to be able to evaluate the relative contributions of motivational and knowledge variables. Most teachers probably assume a motivational explanation of interest effects. If, however, the effect of interest is not so much due to greater effort expended but to greater knowledge, then it would imply that a major barrier to reading comprehension is not the desire to read but a student's limited knowledge of the world.

Some preliminary attempts to separate motivational and knowledge explanations have been made. Asher (1975) created a set of passages, each of which could be made to be about a variety of topics while in fact containing the same vocabulary. Each child received three topics associated with his or her highest rated pictures and three associated with his or her lowest rated pictures. The assignment of topics to passages was random.

The results indicated no difference between high-interest and low-interest performance on these controlled-vocabulary passages. This would tend to support the view that some type of knowledge variable accounts for interest effects when they occur. However, the post-reading rating children made of how much they would like to read more about each passage suggest that the interest manipulation was not very strong. Although children rated the passages associated with high-interest topics significantly higher, the actual difference in ratings was quite small. It seems that in controlling the vocabulary, much of the richness associated with a topic
was lost. The cat paragraph, for example, doesn't seem to have very much to do with cats. Because the manipulation of interest was weak in this study, the comprehension data cannot be confidently interpreted. The controlled vocabulary approach does not seem to be an effective way of testing the motivation versus knowledge explanations.

Two other types of strategies might be attempted in future studies. One approach would be to provide strong external incentive for doing well. If poorer performance on low-interest material results from low motivation, then it should be possible to equalize performance on high- and low-interest material by providing additional incentive to do well. If, however, there are vocabulary or knowledge constraints on low-interest materials, then the additional incentive to try hard should not be that effective in reducing the gap between high- and low-interest performance.

Another strategy would be to pre-assess children's knowledge with respect to certain passages. Perhaps why questions (Bormuth, 1973; Finn, 1973) could be used to assess children's knowledge about specific passages prior to having them read the passages. Perhaps, given enough passages, there would be a sub-set of two or three for each child that would fall into each of the following four cells: high-interest, high knowledge; high-interest, low-knowledge; low-interest, high-knowledge; and low-interest, low-knowledge. This type of experiment could assess the strength of the interest effect, knowledge effect, and the interaction of these two variables.

Another issue that might be explored is whether high-interest material decreases race differences in reading comprehension. The typical findings in previous research is that the gap in reading achievement test performance
widens with each passing year in school (Coleman et al., 1966; Singer, Gerard, and Redfearn, 1975). The interestingness of the material might well have a role to play either because black children are less motivated to read or because they have less familiarity or knowledge about many of the topics they encounter in school.

One interest study to date tested fifth grade black children and white children (Asher, 1976). Each child's interests were first assessed using the picture rating technique and two weeks later all children received a set of six passages from the Britannica Junior Encyclopedia. Three passages corresponded to the child's highly rated topics and three to lowly rated topics. Results indicated that both black children and white children were facilitated by the high-interest material. The strong effect of topic interest for black children is promising in light of the nature of the reading material. It is difficult to conceive of a better representative of standard dialect material than the Encyclopedia Brittanica. The strong effect of interest contrasts with the modest effect that dialect variation seems to have on black children's reading comprehension (Hall and Turner, 1974; Nolen, 1972).

It should be noted, however, that interest and race did not interact in the Asher study. That is, the gap between black and white children's performance remained the same on high- and low-interest material. One possibility is that somewhat easier material would produce different findings. Although the Encyclopedia is said to be for third through sixth grade children (Walsh, 1973), readability analysis using the Dale-Chall formula indicates that the passages are at a seventh grade level (Asher and Markell, 1974). Perhaps if more grade appropriate material had been used, an interaction of
race with interest might have been obtained. It remains to be seen whether high-interest material can reduce the gap between black and white children's reading performance.

One interesting secondary finding from the Asher (1976) study was the pattern of interest ratings for black and white children. There was considerable similarity of interests across race. Johns (1973) also has recently presented data suggesting that black and white children have similar reading preferences. These data suggest that the movement in the 1960's and 1970's to write textbooks with story content for black children may have overestimated the distinctness of black children's interests. Once again, it seems that the assignment of material should be based on individualized assessment of interests.

Another issue deserving attention is whether achievement tests contain considerable amounts of low-interest material. A recent analysis of the characters or actors in achievement test passages indicates that a disproportionate number of characters are male (Faggen-Steckler, McCarthy, and Tittle, 1974). However, as Dwyer (1974) has noted, there is a distinction between sex-role bias and content effects on performance. Story content is probably a more important determinant of performance than sex of the story characters. Research could be done to learn whether children's performance on achievement tests passages is more similar to their performance on low- than high-interest material. If achievement test passages are functionally of low-interest then many children's ability to derive information from reading may be under represented by existing tests.
Summary

Reading is the process of imposing an interpretation on a piece of textual material. The ability to carry out this task efficiently and reliably involves the knowledge of and sensitivity to a variety of linguistic conventions adhered to (implicitly or explicitly) by an author. Difficulty in comprehension can arise whenever the author's assumption about the knowledge, skills, processing capacity, and expectations are unfulfilled. Thus, we see every reading difficulty as a kind of communication failure. We see (at least one of) the goals of reading education to be the collection of skills and knowledge a child brings into the classroom with those additional ones required to meet the assumptions of most authors and the limitations of the written mode of communication. In this report we have proposed to develop methods of isolating and making as explicit as possible those assumptions (and their textual consequences) which appear to cause young children the most difficulty. We suspect that such an analysis can serve at least three purposes: (1) we can inform authors of children's texts about the assumptions that they (the authors) are tacitly making; (2) we can inform teachers of the kinds of assumptions which are typically unfulfilled so that an educational program might be developed in an attempt to supplement a child's knowledge or skills and (3) the patterns of difficulties will serve as a foundation for a more detailed theory of linguistic information processing.


Asher, S. R. Effect of interest in material on sex differences in reading comprehension. Final report to the National Institute of Education (Project No. 3-1324), June, 1975.


Green, C. M. Main clause phenomena in subordinate clauses. Lg, 1976, 52, 382-397.


Kaplan, P. M. Augmented transition networks as psychological models of sentence comprehension. Artificial Intelligence, 1972, 3, 77-100.


Peltz, F. K. The effect upon comprehension of repatterning based on students' writing patterns. Reading Research Quarterly, 1972-74, 9, 603-621.


Thorndike, E. L. The teacher's word book. New York: Teachers College, Columbia University, 1921. (b)


No. 2: Spiro, R. J. Inferential Reconstruction in Memory for Connected Discourse; October 1975. (ERIC Document Reproduction Service No. ED 136 187, 81p., HC-$4.67, MF-$0.83)


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. 7: Ortony, A. Names, Descriptions, and Pragmatics, February 1976. (ERIC Document Reproduction Service No. ED 134 331, 25p., HC-$1.67, 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. 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. 54: Fleisher, L. S., & Jenkins, J. R. Effects of Contextualized and Decontextualized Practice Conditions on Word Recognition, July 1977.


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


No. 61: Spiro, R. J., & Smith, D. Distinguishing Sub-Types of Poor Comprehenders: Overreliance on Conceptual vs. Data-Driven Processes, October 1977.


No. 64: Spiro, R. J., & Martin, J. E. Contextual Factors in the Recall of Alternative Surface Structures, October 1977.