1992

Reader-text match: The interactive effect of reader ability and text difficulty on comprehension monitoring

Kathryn Maelou Baxter
University of Northern Iowa

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9226001

Reader-text match: The interactive effect of reader ability and text difficulty on comprehension monitoring

Baxter, Kathryn Maelou, Ed.D.

University of Northern Iowa, 1992

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

READER-TEXT MATCH: THE INTERACTIVE EFFECT OF READER ABILITY AND TEXT DIFFICULTY ON COMPREHENSION MONITORING

A Dissertation

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

Approved:

Dr. Jeannie L. Steele

Dr. Susann Doody

Dr. Scott Cawelti

Dr. Charles Dedrick

Dr. Jack Yates

Kathryn Maelou Baxter

University of Northern Iowa

May 1992

ACKNOWLEDGEMENTS

This dissertation is dedicated to my mother, who taught me to love learning.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

READER-TEXT MATCH: THE INTERACTIVE EFFECT OF READER ABILITY AND TEXT DIFFICULTY ON COMPREHENSION MONITORING

An Abstract of a Dissertation

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

Approved:

Jeannie H. Steele
Faculty Advisor

John W. Somerville
Dean of the Graduate College

Kathryn Maelou Baxter

University of Northern Iowa

May 1992

ABSTRACT

The purpose of this study was to determine the effect of text difficulty on the comprehension monitoring of above- and below-average readers. In most of the good reader/poor reader comprehension monitoring research, students have been given identical passages. Consequently, the poorer readers must read text that is relatively more difficult for them than it is for the better readers.

In this study, 36 fourth graders, 18 above-average readers and 18 below-average readers, were given text on three levels of difficulty, as determined by the Fry readability formula: (a) a second-grade passage for all students, used to reflect the standard practice of administering the same passages to all subjects; (b) ability-appropriate passages--third-grade for below-average readers and fifth-grade for above-average readers; and (c) difficult passages--fourth-grade for below-average readers and sixth-grade for above-average readers. Students were asked to find anomalous words placed in the materials, a comprehension-monitoring measure known as the Error Detection Paradigm.

It was hypothesized that: (a) better readers would outperform the poor readers overall; (b) there would be an interaction between reader ability and text

difficulty; and (c) text difficulty would affect comprehension monitoring performance. Better readers did outperform the poor readers, but the other two hypotheses were not upheld. The difficulty of the material appeared to have no discernable pattern of effect on comprehension monitoring performance, which was idiosyncratic from reader to reader and from passage to passage.

In an effort to understand the results, other factors influencing the reading process were examined. These factors included such reader factors as reader knowledge of topic and text structure and text factors such as text structure and considerateness. The results of this study brought into question the validity of using a readability formula to establish text difficulty and of using standardized test scores to assess reader ability.

Further research investigating factors influencing the reading process is indicated. Reading must be viewed as a complex, interactive process involving the reader and the text in the construction of meaning.

CHAPTER I

INTRODUCTION AND REVIEW OF LITERATURE

In broad terms, the purpose of this study was to examine the reasons poor readers are poor readers. One possibility is that they lack something that good readers have. Another possibility is that, rather than having a specific deficit, they are simply being asked to read at a level that is not appropriate for them. A large body of research has centered on the first possibility, examining the differences between good readers and poor readers.

Typically, readers are classified as good or poor based on such criteria as standardized test scores or performance relative to grade level. The good reader/poor reader research has attempted to discover what good readers have that poor readers lack. "A deficit in any one area has been assumed to provide an explanation of reading (dis)ability" (Lipson & Wixson, 1986, p. 115). The model of existential proofs (Pearson & Gallagher, 1983) underlies this research. In the existential proofs model, if a variable can be shown to affect reading comprehension, and if it appears to be present to a greater degree in the reading repertoire of better readers than in the repertoire of poorer readers,

then that variable should be taught to poorer readers. In other words, find the missing puzzle piece, fit it in, and the puzzle forms a complete picture or, in this case, a good reader.

Researchers have made an extensive search for these missing puzzle pieces. Poorer readers have been found to be inferior to better readers in many studies in which the two groups were compared on such wide-ranging variables as eye movements (Fairbanks, 1937; Miles & Segel, 1929), ability to form mental images (Gambrell & Bales, 1986), inferencing ability (Davey & Macready, 1985), and knowledge and application of story structure (McConaughy, 1985; Rahman & Bisanz, 1986). However, the variable that has been most extensively studied in the good reader/poor reader research is the ability to monitor one's comprehension.

Comprehension monitoring is an important metacognitive activity that has been characterized as the executive system of the mind (Anderson, 1975), an executive function (Paris, Cross, & Lipson, 1984), or executive control (Garner, 1987). It generally is thought of as having two components or stages, evaluation and regulation (Baker, 1985; Baker & Brown, 1984a; Garner & Kraus, 1981-1982).

Good reader/poor reader studies generally have found that better readers are better monitors of their own

comprehension than poorer readers. In these studies, researchers have attempted to eliminate alternative explanations for their results by controlling for variables such as intelligence (McConaughy, 1985) and decoding ability (August, Flavell, & Clift, 1984). However, even though most have carefully chosen or constructed the reading passages used in their studies by such methods as using readability formulas (Miller, 1987) and standardizing the number of sentences (Winograd & Johnston, 1982), the emphasis has been on the general reading ability of the reader, with little regard for the interactions among individual readers, texts, and contexts (Lipson & Wixson, 1986). Of particular interest here is the interaction between reader and text, which will be referred to in this paper as reader-text match. It is the purpose of this study to investigate the effect of text difficulty on the comprehension monitoring performance of readers of different abilities.

Reader-text match is one of the central precepts of reading instruction. It is an extremely complex question that involves consideration of many factors within the reader and within the text. Recent research and thinking about the nature of reading views the relationship between the reader and the text as equal, dynamic, and interactive (Farris & Triplett, 1986).

Current theories about understanding text emphasize the active role of the reader (Gillet & Temple, 1990). One way of looking at the variables within readers that has led to a new understanding of reading is schema theory (Latvala, 1989). Schemata are regarded as the "building blocks of cognition" (Rumelhart, 1980, p. 33), "a knowledge structure, or framework, which interrelates all of one's knowledge about a given topic" (Richgels, 1982, p. 54), or "an abstract knowledge structure derived from repeated experiences with objects and events" (Garner, 1987, p. 3). A reader's schema or background knowledge about a topic is difficult to assess, but it has a significant impact on his or her comprehension of a particular text.

The most common attempt to assess factors within a reader is the practice of determining reading levels for an individual. Betts (1936) set criteria for three reading levels that may be identified for each reader. The first, the independent level, is the level at which a person is able to read on his or her own. It is a comfortable, easy level for that person. The second, the instructional level, is the level at which the material is appropriate for that person under instructional conditions, as the term implies. Finally, Betts identified the frustration level. This is the level where the material becomes too difficult for a person.

This is the simplest view of personal reading level and is usually assessed through the use of narrative text. Not generally considered in this kind of reading assessment are less easily assessed reader factors, such as the background knowledge or interest of the student as they pertain to a specific passage or the special characteristics of the passage itself, such as concept density or text structure (Wilson & Gambrell, 1988).

Text difficulty, if it is assessed at all, is most typically determined by a simple readability formula that often considers only such factors as word and sentence length. Such text characteristics as concept density, rhetorical patterns, and text considerateness are typically disregarded.

The reading strategies of students vary according to reader-text match. Placed in frustration-level materials, students by definition make many more errors, but they also make different kinds of errors than they do in materials at their instructional level (D'Angelo & Mahlios, 1983; Kibby, 1979; Pikulski & Shanahan, 1982).

Despite what is known about reader-text match, the reading strategies of good readers and poor readers of the same age when they are placed in materials matched to their individual reading levels have not been adequately assessed. In some studies (Erickson, Stahl, & Rinehart, 1985; Paris & Myers, 1981), the difficulty of the text

has been varied, but all students were given the same text regardless of their reading level. A more typical practice has been to give all students material at or below the instructional level of even the lower readers; however, researchers following this practice have not considered the fact that such material was relatively much easier for the better readers.

One attempt to solve the problem of reader-text match is the reading level design, which in a sense brings the reader to the text by comparing younger, adequate readers with older, disabled readers of the same reading ability on their performance in identical materials. Although the reading level design is an ingenious attempt to solve the reader-text match problem, this paradigm has had limited use and is not without problems of its own, such as the higher mental ages and better memories of the older readers (Goswami & Bryant, 1989). It appears that few attempts have been made to solve the reader-text match problem by bringing the text to the reader--that is, by giving good readers and poor readers of the same age different passages that are appropriate for their differing ability.

This chapter will present a brief overview of the research comparing good readers with poor readers. A review of comprehension monitoring studies that have compared good readers and poor readers will follow. Next

will come a discussion of the problems involved with matching readers and texts. The chapter will conclude with a statement of the purpose of the study and the research questions.

Good Reader/Poor Reader Research

A common approach to reading disability research contrasts groups of varying ability. Indeed, Singer (1982) called this the prevailing approach to reading disability research. The paradigm has prevailed so that there exists today a large body of research focusing on the differences between good readers and poor readers.

Good readers and poor readers have been compared on a wide variety of variables and, usually, good readers have outperformed poor readers. Early researchers began by looking at the difference in eye movements in good and poor readers (Anderson, 1937; Fairbanks, 1937; Miles & Segel, 1929). Subsequently, researchers have studied such diverse variables as purposes for reading (Smith, 1967), reading time (Cromer, 1970), task-attending behaviors (Gambrell, Wilson, & Gantt, 1981), use of graphic information and context (Biemiller, 1979), word prediction (Allington & Strange, 1978), lexical access (Briggs & Underwood, 1987), and reasoning strategies (Kavale & Schreiner, 1979). Also considered have been inferencing ability (Bridge, Tierney, & Cera, 1978; Davey

& Macready, 1985), ability to form mental images (Gambrell & Bales, 1986), and story structure and schemata (McConaughy, 1985; Rahman & Bisanz, 1986). There have been a few studies in each of many areas. In one area, however, there is a much larger body of research. These are the studies that compare good readers and poor readers on their ability to monitor their own comprehension (Erickson et al., 1985; Garner & Taylor, 1982; Markman, 1979; Paris & Myers, 1981).

Comprehension Monitoring

A variable frequently studied by researchers in reading is comprehension monitoring, an aspect of metacognition. Metacognition has been informally defined as thinking about thinking (Babbs & Moe, 1983) or cognition about cognition (Flavell, 1981). More formal definitions indicate that metacognition refers to the deliberate, conscious control of one's own cognitive actions (Baker & Anderson, 1982; Brown, 1980; Garner, 1987), and it appears to have grown out of the concept of metamemory (Brown, 1975; Flavell, 1979). Related terms are cognitive monitoring, comprehension monitoring, and metacomprehension.

There seems to be much confusion about the use and definition of these terms and their relationship to each other. Baker and Brown (1984b) clearly described their hierarchical relationship. In their outline,

metacognition is composed of two parts: awareness and regulation of one's comprehension. The regulation aspect, also known as cognitive monitoring, includes checking, planning, evaluating, testing, revising, and remediating. Baker and Brown stated that a large part of this cognitive monitoring is comprehension monitoring, sometimes called metacognition (Fitzgerald, 1983; Raphael, Myers, Tirre, Freebody, & Fritz, 1981). Comprehension monitoring has three aspects: (a) keeping track of the success of one's comprehension, (b) ensuring that success continues, and (c) remediating, if necessary. It is more commonly discussed as having two components: evaluation and regulation (Baker, 1985).

In the evaluation stage, a reader becomes aware that comprehension has broken down. In the regulation stage, the reader repairs or "fixes up" comprehension by using such debugging strategies (Wong, 1986) as rereading, scanning ahead, ignoring confusions, forming hypotheses, and changing reading rate (Pitts, 1983).

The two components of comprehension monitoring are not inseparable. It is incorrect to assume that a reader who is able to evaluate his or her comprehension and becomes aware of a breakdown will necessarily be able to follow through with the regulation phase and repair that breakdown (Baker & Brown, 1984b). Comprehension monitoring research began with an oral paradigm (Markman,

1977), but there has been more comprehensive investigation of comprehension monitoring in reading than in listening. Comprehension monitoring most commonly is viewed as applying mainly to reading comprehension of connected discourse (Wagoner, 1983).

One reason comprehension monitoring has been so widely studied is that it is accepted as an important factor in reading comprehension. "It has long been assumed that comprehension monitoring is crucial to effective reading" (Baker & Anderson, 1982, p. 282). Winograd and Johnston (1982) insisted that, "The definition of fluent reader must include a reference to the ability to self-check and self-correct reading strategies" (p. 4). Paris, Lipson, and Wixson (1983) called comprehension monitoring "a key to reading proficiency (p. 300), and Zabrucky, Moore, and Ratner (1985) called it "a critical skill in reading comprehension" (p. 241). The primary value of comprehension monitoring is that it allows readers to take responsibility for their own learning and comprehension (Baker, 1987; Wilson & Gambrell, 1988).

The acceptance of comprehension monitoring's crucial role in reading comprehension is evidenced by the inclusion of comprehension monitoring in recent texts on the teaching of reading (Durkin, 1989; Gillet & Temple,

1990; Lapp, Flood, & Farnan, 1989), by the proliferation of training studies wherein students are trained to use comprehension monitoring (Gambrell & Bales, 1986; Grabe & Mann, 1984; Guthrie, 1983; Miller, 1987; Reis & Spekman, 1983; Short & Ryan, 1984; Siedow & Fox, 1984; Tregaskes & Daines, 1989; Winograd & Johnston, 1982), and by the abundance of articles on how to improve students' comprehension monitoring through instruction (Babbs, 1984; Babbs & Moe, 1983; Bransford, Vye, & Stein, 1984; Davey, 1983; Fitzgerald, 1983; Hahn, 1984; Lapp & Flood, 1984; Palinscar & Brown, 1986; Paris et al., 1983; Pitts, 1983; Ryan, 1981).

Since good readers usually have been found to be better monitors of their comprehension than are poor readers (Garner & Taylor, 1982; Winograd & Johnston, 1982), investigators set out to find ways of training readers to monitor their own comprehension. The training studies described below are based on Pearson and Gallagher's (1983) model of existential proofs. They were attempts to see if the missing "puzzle pieces" that supposedly make up good readers can be taught to readers to improve their comprehension.

Training Studies in Comprehension Monitoring

Once it is assumed that good readers do something that poor readers don't do, the obvious next step is to see if it is possible to train students to emulate the

strategies of good readers. In the studies reviewed here, the primary strategy to be emulated is comprehension monitoring. In these training studies, all students, regardless of ability, received the same passages in both training and evaluation of the effect of the training.

In 1982, Winograd and Johnston hypothesized that poor readers' comprehension monitoring performance would improve if they were given instruction in schema selection. The subjects were sixth-grade students who were classified as skilled and less skilled on the basis of their scores on a standardized reading test. The researchers found that the schema selection treatment had no significant effect.

Reis and Spekman (1983) conducted a study to see if direct instruction in performing an editing function would improve the ability of sixth- and seventh-grade poor comprehenders to detect two kinds of inconsistencies, text-based and reader-based. Following training, the group receiving training scored significantly higher on reader-based inconsistencies than the control group, but there was no significant group difference for text-based inconsistencies.

The ability to identify inconsistent information was also the focus of a training study conducted by Grabe and Mann (1984). Their subjects were good and poor readers

among fourth and fifth graders and among college students. The students played reading games in which the task was to identify consistent and inconsistent statements. The training resulted in a significant increase in both good and poor readers' ability to identify consistent statements but not inconsistent statements.

Two training studies attempted to improve comprehension monitoring performance through instruction in text structure. Siedow and Fox (1984) investigated the effects of instruction in expository material top-level structures such as cause/effect and comparison. Their subjects were good and poor tenth-grade readers. They found that their instruction significantly improved poor readers' ability to use top-level structure, but that it had no significant effect on good readers' performance. Short and Ryan (1984) investigated the effects of training in story grammar and in attribution training on male fourth-grade poor readers. Attribution refers to knowing why we know or fail to know or to accomplish something (Guthrie, 1983). In this study, subjects were trained to recite such attribution self-statements as, "Try hard" and "Praise yourself for a job well done" before reading a story. For the story grammar training, subjects were taught to ask themselves such questions as, "Who is the main character?" and "How

did the story end?" Story grammar training had a greater impact on the performance of the poor readers than did the attribution training. Short and Ryan found that the story grammar training brought the poor readers up to the level of the more skilled readers who were used as a comparison group in the posttest only.

Miller (1987) investigated the effect of self-instruction on the ability of average and above-average fifth-grade readers to detect errors. She found that the above-average readers benefited more from the self-instruction than did the average readers.

Mental imagery has been considered as a strategy that may improve students' comprehension monitoring. Gambrell and Bales (1986) trained fourth- and fifth-grade poor readers in the use of mental imagery. The posttest error detection performance of the training group was significantly better than that of the control group. Tregaskes and Daines (1989) used mental imagery training along with training in four other metacognitive strategies: summary sentences, webbing, self-questioning, and use of monitoring or "click" cards. The subjects in this study were sixth graders of varying abilities. The results showed a significant difference favoring the training group over the control group.

The subject of comprehension monitoring has gone through the natural progression from investigation of

what it is that good readers do that poor readers do not do, to training studies to see if students' comprehension monitoring can be improved, and into instructional practice in the schools and instructional recommendations in the professional literature. The findings from research have passed into practice. However, the comprehension monitoring research, like other good reader/poor reader research, was largely conducted without consideration for the match between individual readers and specific texts.

The body of research on which instruction in, and assumptions about, comprehension monitoring is based began with Markman's (1979) study of children's ability to detect inconsistencies when they are listening. However, this discussion will be primarily limited to comprehension monitoring in reading and to studies using elementary and secondary students as subjects.

Good Reader/Poor Reader Studies in Comprehension Monitoring

Much, but not all, of the research in comprehension monitoring has been conducted using reading ability as an independent variable, comparing readers of differing abilities, usually good readers and poor readers. There is a great variation in the measures that have been used to assess comprehension monitoring.

In most of the comprehension monitoring studies, some effort was made to control the characteristics of the passages used. Many of the selections were meticulously constructed by the researchers (August et al., 1984; Garner & Kraus, 1981-1982; Zabrucky et al., 1985). Others were carefully selected from published materials (Erickson et al., 1985; Paris & Myers, 1981; Stahl, Rinehart, & Erickson, 1985-1986). Sometimes a further control was introduced by checking the students on their recognition of the words in the passages (August et al., 1984; Winograd & Johnston, 1982) or by asking teachers if all the passages could be read by all the students (Owings, Peterson, Bransford, Morris, & Stein, 1980). In an effort to be "fair," the same passages were given to all of the readers regardless of their reading ability. However, what appears to be "fair" may not be fair at all. In attempting to control the variables in their studies by giving all students the same passages, researchers have given poorer readers materials that are relatively more difficult, sometimes much more difficult, for them than they are for the better readers. This section will discuss the ways in which comprehension monitoring has been assessed in studies using groups of readers of differing abilities. Particular attention will be given to the question of the level of text

difficulty relative to reader ability, or reader-text match.

Text lookbacks. Text lookbacks have been used as one measure of comprehension monitoring. The assumption is that if readers refer to previously read text, they do so because they have become aware of experiencing some sort of comprehension breakdown (evaluation), and they are taking measures to repair that breakdown (regulation).

Garner and Reis (1981) compared the frequency of text lookbacks in sixth-, seventh-, and eighth-grade students who were rated as good and poor comprehenders by their teachers. This study did not investigate students' lookback behavior during the reading of the text. Rather, it examined lookbacks while students were answering postreading questions. All of the students were given the same passage. No reading level was specified. Garner and Reis found that poor readers less frequently referred back to the text in answering a question, but that very few students of any ability used lookbacks at all.

In August et al.'s (1984) study, standardized test scores were used to designate fifth-grade good and poor readers who were compared in their use of lookbacks during reading. The researchers deleted sections of the stories so that the remaining information was

inconsistent. The students read the stories on a computer so that it was quite clear whether or not they used lookbacks. All students received the same second-grade passages. A readability formula based on sentence and word length was used to rate the passages. The results of the lookback measure were not analyzed statistically because of the low number of students of either ability level who used lookbacks.

Text lookbacks were used as one of the measures of good and poor sixth-grade readers' evaluation and regulation of their comprehension in Zabrucky and Ratner's (1986) study. Good readers were found to look back at inconsistencies more often than poor readers. They also gave more accurate verbal reports of passage consistency and were better able to recall text inconsistencies. However, children in both groups read inconsistencies more slowly than consistent information. The passages used in this study had a third-grade Fry readability. All students received the same passages.

Think-aloud protocol analysis. Protocol analysis studies collect self-report data from subjects as they read. Subjects are encouraged to think aloud, and transcripts are made of their comments. These transcripts are then analyzed for number and nature of comments that might indicate the subjects' use of comprehension monitoring.

Smith (1967) conducted an early comprehension monitoring study using protocol analysis before the term comprehension monitoring came into use. She concluded that responses by good readers demonstrated more awareness of reading processes and greater use of strategies such as rereading and reviewing content. Her subjects were twelfth graders who were designated good or poor readers on the basis of a standardized test. The reading selection used for all students in the study was seventh- to eighth-grade level in reading difficulty.

A decade later, Olshavsky (1976-1977, 1978) conducted two protocol analysis studies that clearly assessed comprehension monitoring, but again, the term itself was not used. In her 1978 study, which Olshavsky described as "exploratory," she examined the strategies used by good and poor 11th-grade readers, who had been chosen on the basis of standardized test scores, to accommodate increasingly difficult material. All students were given the same four passages, 7-8, 9-10, 11-12, and 13-15 grade levels, as measured by the Dale-Chall (1948) readability formula. She found that there was no significant difference between the two groups of readers in the use of such strategies as forming hypotheses and substituting synonyms, and that both groups used fewer strategies as the difficulty of the material increased.

In the other study, Olshavsky (1976-1977) used protocol analysis to investigate strategy usage by 10th-grade good and poor readers. The materials were on a 10th-grade reading level as determined by the Dale-Chall (1948) formula. Olshavsky found that good and poor readers generally use the same kinds of strategies, but that good readers use them more frequently.

Interviews. Another kind of self-report data that has been collected is from interviews with subjects about their awareness of their own metacognitive strategies. Interviewers look for "meaning-oriented" responses in these interviews as an indication that subjects are using comprehension monitoring.

In one such study, good and poor comprehenders who were selected by their seventh-grade teachers were interviewed about their approach to reading comprehension in a study by Garner and Kraus (1981-1982). They used six interview questions such as, "What things does a person have to do to be a good reader?" and "What do you do if you don't understand something you are reading?" (p. 7). They scored the interview protocols by assigning one point for any response that mentioned anything related to meaning, comprehension, understanding, or making sense. All other responses were scored zero. They found that good readers gave meaning-related responses significantly more often than poor readers.

Erickson et al. (1985) also used the six-question interview and scoring system designed by Garner and Kraus. Their findings contrasted with those of Garner and Kraus (1981-1982). They found no significant differences in number of "meaning" responses between above- and below-average sixth-grade readers who had been identified through standardized test scores. All students were given both third- and sixth-grade passages from Reader's Digest Skill Builders (1960).

Difficulty and comprehensibility ratings. Another subjective measure of comprehension monitoring is asking subjects to assign difficulty, comprehensibility, or liking ratings to passages. The idea is that good comprehension monitors will find incomprehensible or inconsistent passages more difficult and less understandable. Differences in the students' perception of difficulty would not be caused by the difficulty of the material because of its reading level, because in these studies all of the students received identical passages regardless of their reading ability.

Garner (1980, 1981) conducted two studies that incorporated a difficulty rating as one of the measures. The students, regardless of ability, were all given the same six passages. No readability was specified. The difference in difficulty was presumed to come from the presence of errors. The procedure for taking the

difficulty rating was very simple. Subjects, who were good and poor seventh- and eighth-grade readers in the first study and all poor fifth- and sixth-grade readers in the other, were asked to check one of three boxes at the bottom of each passage they read. The choices were, "This part was very easy to understand," "This part was ok," and "This part was difficult to understand." In the first study, good readers demonstrated an awareness of actual difficulty by rating nearly all consistent-information segments of the passages as easy to understand, and inconsistent information segments as being more difficult. Poor readers did not make this distinction. The researchers concluded that good readers noticed the text disruptions, but poor readers did not. The second study confirmed the results of the first. The poor comprehenders reported little difference in the difficulty of the consistent and inconsistent passages.

In a 1980 study, Owings et al. used researcher-constructed passages of no specified difficulty for both the good comprehenders and the poor comprehenders. Fifth-grade students who were considered good comprehenders because of their standardized test scores were better able to distinguish between difficult and less difficult stories than were poor comprehenders. The "difficult" stories had been written so that they did not

make sense. The good comprehenders were also able to explain why they rated the stories as they did.

A difficulty rating was also used by Erickson et al. (1985) in a study of comprehension monitoring in sixth graders who scored above- and below-grade level on a standardized test. One of the passages given to all students was at the third-grade level and one at the sixth-grade level. They found no significant difference between the two groups.

Zabrucky et al. (1985) used a comprehensibility rating as a measure of comprehension monitoring. The passages used for all students in the study were constructed by the researchers. No reading level was specified. In addition to the comprehensibility rating, they asked the good and poor second-grade readers who had been selected through the SRA assessment device to say whether they liked the stories or not. The researchers predicted that both comprehensibility rating and liking would be related to error detection for good readers but not for poor readers. Their data supported that prediction. They interpreted that finding to mean that the good readers took a meaning-based approach to reading while the poor readers did not.

Garner and Taylor (1982) included a comprehensibility rating in their study of comprehension monitoring in good and poor comprehenders in Grades 4, 6,

and 8. All readers received the same researcher-constructed passages of no specified readability, each of which contained a major informational inconsistency. They used the rating as the basis for questioning the readers about why they found the passage easy or difficult to understand. The same passages were used by Garner and Kraus (1981-1982) with seventh graders who were labeled by their teachers as good comprehenders and poor comprehenders. In both studies, the good readers were better able to choose the more understandable passages and to explain why they were more understandable.

Reading time. Reading time has also been used as a measure of comprehension monitoring. The assumption made by researchers who select this measure is that students who are monitoring their comprehension will regulate their reading by varying their rate when reading inconsistent information. In other words, they will read inconsistent material more slowly than they do consistent material.

Owings et al. (1980) conducted a study using fifth graders who were judged as successful or less successful based on their standardized test scores. No readability was specified for the passages, which were, as usual, given to all the students. Among other measures, they used reading time. They found that easy stories (no

errors) were read faster than difficult stories (included errors) for both groups of readers, and that more successful readers read everything faster than the less successful readers. They did not, however, analyze the difference score for each group of readers, so this information is limited in its use.

August et al. (1984) used reading time as one of the measures in their study comparing skilled and less skilled fifth-grade readers. All students received the same second-grade passages. They found that the skilled readers spent significantly more time on the inconsistent stories than the less skilled readers.

Analysis of oral reading. One of the most widely accepted measures of comprehension monitoring is a student's oral reading. Of particular interest are self-correction of errors and semantic acceptability of miscues.

Beebe's (1980) study is of particular interest because it indicated the validity of using oral reading analysis to give information about both oral and silent comprehension monitoring. The subjects were fourth-grade boys who read either at or above grade level. She found self-corrections of substitution errors to be positively correlated with both a conventional comprehension measure and a retelling.

Paris and Myers (1981) used material on two reading levels, third- and fifth-grade, which contained both nonsense words and nonsense phrases to investigate the comprehension monitoring of good and poor readers among fourth graders. A commercial reading assessment survey was used to select the readers. All students received the same materials, which were taken from the Spache Diagnostic Reading Scales (Spache, 1953). The researchers recorded repetitions, hesitations, and self-corrections in the subjects' oral reading. After the reading, students were directed to underline anomalous information and were asked comprehension questions. Finally, a free recall was elicited for each passage. Both the design of and the results from this study were so complicated that it is difficult to draw any conclusions from it. For example, poor readers noticed more nonsense words in the third-grade stories than did the good readers, but fewer nonsense phrases.

Oral miscues were also analyzed by Leslie (1980), who noted whether miscues were syntactically and semantically appropriate. Standardized test scores were used to select the subjects, who were second-grade average readers and below-average readers in third to sixth grade. The materials used for all students were of a 2.1 or 2.2 reading level. Results showed that below-average readers made more miscues that were

graphically similar to the word than average readers and that their miscues tended to be less often semantically acceptable.

Cloze tasks. The use of context is closely associated with comprehension monitoring. One measure of use of context is an adaptation of the cloze technique, developed by Taylor (1953). Originally, the cloze procedure was developed to establish students' reading levels and to match them with materials. Subsequently, it has also been used to assess students' comprehension. In a cloze passage, words are deleted from the text and students are asked to fill them in. Typically, the first and last sentence is left intact, and every fifth word is deleted in the remainder of the passage.

DiVesta, Hayward, and Orlando (1979) used cloze passages to assess the comprehension monitoring proficiencies of good and poor sixth, seventh, and eighth graders, who were selected from their standardized test scores. The passages used for all students were within the sixth- to eighth-grade range of reading level (Fry, 1968). Within the cloze task there were two conditions, running text (RT) and subsequent text (ST). In the first condition, blanks could be filled in by using information already found in the story. In the subsequent text condition, it was necessary to read ahead and then go back and fill in the blanks. As expected, the poor

readers were outperformed by the good readers. Also, the difference between the two scores (RT and ST) was twice as great for poor readers as for good readers. In other words, poor readers did not seek out information in subsequent parts of the text to help them fill in the blanks.

Error detection. Many comprehension monitoring studies have used some variation of error detection to measure comprehension monitoring. In what has come to be called the Error Detection Paradigm (Winograd & Johnston, 1982), errors of various types are introduced into text and students are asked to find them. The assumption is that if students are monitoring for comprehension, they will perform well on this task. "Error detection is often considered to be an index of comprehension monitoring that also serves as a measure of comprehension" (Paris et al., 1984, p. 1243). Several error detection studies have used an editor persona, asking students to act as editors to detect problems (Garner & Kraus, 1981-1982; Garner & Taylor, 1982) or as consultants (Winograd & Johnston, 1982).

Baker (1985) has identified three basic standards for evaluating text comprehension, all of which have been tested in the comprehension monitoring studies. The first is the lexical, or word standard, which requires readers to determine whether a word makes sense in a

passage. This standard has been tested through the introduction of anomalous or nonsense words in passages. Next is the syntactic standard, which is tested by disrupting the grammatical construction of a passage. Finally, there is the semantic standard, which requires the reader to evaluate the meaning of the text as a whole. This standard has been tested by disrupting the meaning of passages through such devices as omitting text or using statements contrary to the prior knowledge of the readers.

Directed underlining of nonsense words and phrases was the error detection task in a study conducted by Paris and Myers (1981). The students were told to underline any words or sentences in the story that they didn't understand. The poor readers noticed more nonsense words in the stories written at their own third-grade reading level than did the better readers, but their performance was considerably inferior to that of good readers on the fifth-grade passages. The surprising finding was that neither group of readers did very well at marking nonsense words and phrases.

Indeed, one concern about the Error Detection Paradigm has been the nondetection of errors. Garner (1987) estimated that one-third to almost two-thirds of intentionally inserted errors went reported in the comprehension monitoring studies she reviewed. In many

studies (August et al., 1984; Garner & Taylor, 1982; Winograd & Johnston 1982), a remarkable number of errors went undetected. This has led to the recommendation that if errors are to be detected, they must be quite blatant in nature.

Winograd and Johnston (1982) used contextually anomalous sentences as their error detection task. The sixth graders were asked to act as consultants in determining the comprehensibility of passages that they were told had been written by other sixth graders. Actually, the passages had been written by the researchers. No reading level was specified. Students' detection of errors was assessed through an interview concerning the comprehensibility of the materials. Winograd and Johnston found, as did Paris and Myers (1981), that although good readers did significantly better than did poor readers, neither group did particularly well.

Garner's name is perhaps most strongly associated with the Error Detection Paradigm because of the number of studies she conducted using the procedure. Garner (1981) used material containing two error types. Inconsistent information was introduced into the researcher-constructed passages received by all readers as were polysyllabic words that were judged to be too difficult for students of that age. As in the Winograd

and Johnston (1982) study, error detection was assessed through an interview that included such questions as, "Can you tell me what made this passage difficult?" Garner found that, while the poor readers identified difficult words as sources of difficulty, they were unable to detect textual inconsistencies.

Garner (1980) and Garner and Taylor (1982) used the same passages containing inconsistencies involving number changes for all students. The two studies were similar, but in the Garner and Taylor study more attentional assistance was given to those students who did not at first detect the errors. Good readers performed better than poor readers in both studies. Also, good readers benefited more from the extra help provided in the second study than did the poor readers.

In a study conducted by Garner and Kraus (1981-1982), there were again two types of errors. In this case, however, one was between sentences and one was within sentences. All students received the same passages. No reading level was specified. The researchers found that none of the poor readers could detect any of the inconsistencies. The good readers were somewhat successful at finding errors between sentences and very successful at finding errors within sentences.

The two types of errors selected by Zabrucky et al. (1985) were close and far inconsistencies in the passages

they had constructed. No readability was specified. In the close condition, the contradictory sentences were adjacent. In the far condition, they were separated by several other sentences. The second graders in the study were told that the researchers needed help in determining if the stories would be good stories for other students their age to read and that two of the four stories they would read were "silly." They were asked to underline the sentences that said the opposite of an earlier occurring sentence in the "silly" stories. Students were also asked how well they liked each story and how understandable they thought it was. The focus of the study was to determine whether performance (error detection) and verbal (liking and understanding) measures were correlated. They found that they were related for the good readers but not for the poor readers.

In a more recent study conducted by Zabrucky and Moore (1989), error detection was used in the pilot test of their passages. They wanted to ensure that the subjects had the ability to detect the inconsistencies. In the pilot test, the students were asked to find the problems in the text. They found that the students could easily detect the inconsistencies.

August et al. (1984) used computers in their study in which the errors were in the form of missing information. The fifth graders were told that some

stories might have a page (computer screen) missing. If the students said that there was indeed a page missing, they were handed the story on index cards, one computer screen per card, and asked where they thought the missing page should go. Significant group differences favoring the good readers were found for both detection of missing pages and correct placement. However, all students' comprehension of what they had read was assessed by asking them to give a gist recall of each story, and there was no significant difference between the groups on this measure.

Two error detection studies, Stahl et al. (1985-1986) and Erickson et al. (1985), used conceptual tempo (impulsive/reflective) as an additional independent variable. The investigators hypothesized that the reflective students would perform better than the impulsive students. Both studies were conducted with above- and below-average sixth-grade readers.

In the first study, the error types were textual inconsistencies and factual inconsistencies (Stahl et al., 1985-1986). The materials were written at a fourth-grade reading level, chosen because that level was below the tested reading level of nearly all the students. Predictably, the total group of above- average readers out-performed the below-average readers on the error detection task. However, below-average reflective

readers did nearly as well on detection of textual inconsistencies as did reflective above-average readers, and there was no significant difference in the performance of impulsive below- and above-average readers in their detection of either kind of error.

In the second study, the error types selected were nonsense words and rearrangement of sentences (Erickson et al., 1985). Passage difficulty was included as an independent variable. All students received the same third-grade and sixth-grade passages. As in most other good reader/poor reader studies, there was a significant difference favoring the above average readers on the error detection task. However, the researchers were again surprised by the results of the conceptual tempo investigation. They found that, contrary to their expectations, the impulsive below- average readers performed nearly as well as did both of the above-average groups, whereas the reflective below- average readers did significantly worse. In examining the differences introduced by varying the reading levels of the materials, the researchers found that, although all readers performed similarly on the lower level passage, the below-average readers did significantly worse on the more difficult passage than did the above-average readers.

This difference in performance according to the difficulty of the materials should have been no surprise to the researchers. It is reasonable to assume that they and all the other researchers mentioned previously were familiar with the concept of individual reading levels and the differences in students' reading at each of the levels. Despite the fact that the match of the reader with the text is one of the central precepts of reading instruction, it has been given almost no attention in the good reader/poor reader research. The following section will discuss the problems involved in matching readers and text.

Reader-Text Match

Given the degree of acceptance of the concept of matching the ability of readers with the difficulty of texts, it would be expected that the researchers who have conducted comprehension monitoring studies would have had some concerns about the fit of the text and the reader. This is true on a very limited basis. Most of those who have expressed concern have done so indirectly.

Concerns About Reader-Text Match

Researchers have made attempts to control task variables by training students on the tasks they are to perform. They have tried to standardize testing conditions and so remove the context confound. They have even occasionally tried to account for differences in

background knowledge. What has seldom been done, except through the use of the reading level design that uses students of different ages, is to attempt to match the ability of the reader with the difficulty of the text.

A writer who did express her concern directly was Wagoner (1983). In her review of the research on comprehension monitoring, she described the Paris and Myers (1981) study in which fourth-grade subjects were given passages on two levels. Wagoner says, "A serious problem with this study was that the reading levels of the materials (third- and fifth-grade levels) were more difficult for poor readers than for good readers (average achievement 2.8 and 5.4, respectively), thus confounding poor readers' difficulty-awareness with their problem-awareness" (p. 336). It is puzzling that Wagoner made this comment only about the Paris and Myers study. The same would be true of any of the comprehension monitoring studies whether they used one level of material or more than one. Nevertheless, Wagoner did perceive the problem.

Several researchers have expressed concern about the problems presented by experimental materials. Garner and Anderson (1982) indicated that the materials used in comprehension monitoring studies have an important influence on performance in error detection tasks. They discussed difficulties such as differences in background

knowledge and the magnitude of the inconsistency introduced by the researchers, as well as discussing their attempts to standardize the passages in terms of length, error placement, and difficulty of vocabulary. One factor they failed to recognize is reading ability. That is, that in their study, as in many others, the "standardized" passages were relatively more difficult for the poor readers in their study than for the good readers.

Raphael et al. (1981) pointed out the need for a methodology that would identify points at which material ceases to be comprehensible for readers. They discussed the effects of word difficulty, topic familiarity, and structure on subjects' performance.

Pitts (1983) said, "Whether one comprehends at the subconscious or conscious level depends on characteristics of both the reader and the text" (p. 517). He mentioned cognitive styles, purposes for reading, and familiarity with the information in the text as factors that may influence the readers' performance. Once more, the much simpler question of simple reading level vs. reading ability was not discussed.

More recently, Brown, Armbruster, and Baker (1986), writing about metacognition in reading and studying, commented on the difficulty of matching texts and readers. They considered vocabulary, syntax, clarity of

presentation, structure, and topic as factors that contribute to text difficulty.

Concern has also been expressed about the problem of reader-text match in the literature on miscue analysis or oral reading analysis. Leu (1981-1982) discussed the problems presented by relative passage difficulty. He explained that if two groups of readers of different reading proficiencies read the same passage, the passage will be relatively more difficult for one group than for the other. He hypothesized that the less proficient readers are probably forced into using a processing strategy that focuses more attention on graphic information and less on contextual information. If this is correct, it would account for the often-stated "fact" that poor readers read on a word level whereas good readers read on a whole-text level.

It appears that although many educators recognize the problem of fitting reader ability to text difficulty, only a few researchers have included differential text difficulty in their studies. Two previously discussed studies that have included a text difficulty factor are Paris and Myers (1981) and Erickson et al. (1985).

There was a very low percentage of error detection by both good and poor readers in Paris and Myers' (1981) study. The results, however, were interesting for another reason. The poor readers performed better than

the good readers on detecting anomalous words in the third-grade story, although they were not as good as the better readers at spontaneous monitoring of incomprehensible text when the material difficulty level was fifth-grade. This appears to indicate that poorer readers do monitor their reading, but that text difficulty affects their ability to do so.

Erickson et al. (1985) used a third-grade passage and a sixth-grade passage for all readers in their comprehension monitoring study, which included conceptual tempo (impulsive/reflective) as an independent variable. They found that the above-average readers and the below-average readers performed similarly on the third-grade passage, but that the above-average readers did better than the below-average readers on the sixth-grade passage, although they did not do as well as they had on the third-grade passage. They concluded, "error detection appears not to be a monolithic ability; there are complex interactions between reader characteristics, text characteristics, and error detection types" (p. 250). This confirmed Paris and Myers' (1981) finding that text difficulty affects the ability to monitor comprehension.

A recent study (Latvala, 1989), involving only average fifth-grade readers, contributed further to the

understanding of the relationship of text difficulty and comprehension monitoring. Narrative passages from the Random House Achievement Program in Comprehension (1988) were used to assess students' ability to detect the two internally inconsistent sentences that had been embedded in the third-grade, fifth-grade, and seventh-grade passages rated by the Fry readability formula (Fry, 1968). Error detection results indicated a significantly greater frequency of errors detected at the third-grade reading level than at the fifth-grade level and at the fifth-grade level than at the seventh-grade level. These results support the conclusions of Erickson et al. (1985) and of Paris and Myers (1981) that text difficulty does affect comprehension monitoring performance.

Although it involved high school students, not elementary students, an even more recent study (Kletzien, 1991) did use good and poor comprehenders and did vary text difficulty according to reader ability. The students were asked to complete a cloze exercise and then were interviewed to determine what strategies they had used to select words to fit the blanks. Amount of strategy use was also compared. Kletzien found that while all subjects used the same strategies on the easy passage, total strategy use declined for poor comprehenders as text difficulty increased.

How, then, is the problem of reader-text match to be solved? One attempt to solve it has been the use of reading level designs. Reading level designs have been used primarily by researchers in the psychology of reading.

In a reading level design, sometimes called a matched reading level design, older, reading-disabled children are compared with younger, average readers. The two groups are matched on reading level (Backman, Mamen, & Ferguson, 1984; Davey, 1989; Stanovich, 1986). The designation commonly used is RL design as opposed to a chronological age (CA) design, which is the one commonly used in good reader/poor reader research, where good readers and poor readers of the same age are compared. Advocates of the RL design argue that it minimizes the problem presented by the relative difficulty of the materials in a CA design. If no significant differences are found between the groups on the variables measured, then it can be assumed that reading-disabled children are not qualitatively different from average readers, but that they are simply delayed in their reading development (Backman et al., 1984).

Although Leslie did not call her 1980 investigation of use of graphic and contextual information by average and below-average readers a reading level study, it fit that paradigm. Leslie's subjects were average-reading

second-graders and below-average readers in third- through sixth-grade. Oral miscues were used as the measure of comprehension monitoring. The results showed that the below-average readers made more miscues that changed the author's meaning while average readers' miscues did not alter meaning.

The reading level design has been used to compare students on a number of variables. Stanovich, Nathan, and Vala-Rossi (1986) compared skilled third-grade readers with less skilled fifth-grade readers on general name-retrieval speed and phonological awareness tasks. They found the performance of the two groups to be remarkably similar. Treiman and Hirsch-Pasek (1985) matched younger average children and older dyslexic children on reading level and compared their accuracy in identifying phonologically regular words. There was no difference in the performance of the two groups on this task.

Most of the studies using the reading level design have been in the area of psychology of reading. However, more recently, Davey and her fellow researchers have used the reading level design in "mainstream" reading research to examine the features of test questions and how they influence the responses of good and poor readers (Davey, 1987, 1988a, 1988b; Davey & LaSasso, 1984; Davey & Macready, 1985). Such variables as testing time,

question format, and lookback testing were examined in this series of studies. The researchers found that there were indeed some differences in the features of test questions that affected good readers and poor readers differently, but others that did not. Davey (1989) stated her rationale for using the reading level design: "Using this approach, any observed differences between reader groups could be more reliably attributed to characteristics of the particular readers rather than to differential task or text difficulty" (p. 695).

The reading level design has not been accepted without question. Several writers have expressed differences about its interpretation (Bryant & Goswami, 1986; Goswami & Bryant, 1989; Mamen, Ferguson, & Backman, 1986). However, the reading level design does represent an attempt to solve the problem of the relative difficulty of materials used in reading disability research by "bringing the reader to the text." It does so by matching of students of varying ages who have the same reading level. However, few attempts have been made to solve the problem by "bringing the text to the reader."

Interactive Nature of Reading

Most contemporary theories of reading regard it as a dynamic, interactive, complex process rather than a static, one-dimensional phenomenon where there is one

simple answer to each question. The term interactive is closely associated with the work of Rumelhart (1980) whose work has served as a basis for the discussion of the interactive nature of reading in recent years (Lipson & Wixson, 1986). This interactive nature has been described in various ways by various authors. One theory views the process of reading as involving the interaction of task complexity and material complexity with the level of development of the individual (Mosenthal, 1984). Farris and Triplett (1986) reported that current theory views the relationship between the reader and the text as equal, dynamic, and interactive. Garner (1987) mentioned three sets of factors that act and interact in reading: knowledge of oneself as a reader, the demands of reading tasks, and the strategies one can employ in reading activity.

Within-reader factors often mentioned include prior knowledge or schema (Gillet & Temple, 1990; Richgels, 1982; Wilson & Gambrell, 1988), attribution (Garner, 1990; Guthrie, 1983), and motivation and interest (Johnston, 1983). Some of the factors to be considered with the text include text structure (Garner, 1990; Gillet & Temple, 1990) and text considerateness, which can itself be based on a multiplicity of factors (Garner, 1987). Some educators would include task demands (Flavell, 1979) and/or contextual factors (Lipson &

Wixson, 1986). "This interactive view considers reading ability to be a relative concept, not a static state" (Lipson & Wixson, 1986, p. 115). Consequently, readers' performances will vary according to type of discourse, quantity and quality of schema or prior knowledge about the topic of a passage, considerateness of the text, knowledge and use of metacognitive strategies, and understanding of the task. Lipson and Wixson made the important point that reading performance "varies for both able and disabled readers as a function of the conditions of the reading situation. This calls into question any model of (dis)ability that seeks to identify only causal factors within the reader" (p. 116).

The reader-text relationship is complex, dynamic, and interactive. Both the text and the reader, as well as the interaction between the two, must be considered in the difficult matter of matching readers and texts.

Determining Text Difficulty

The most familiar way of assigning reading levels to text is through the use of any one of many readability formulas. Even though these formulas fail to take into consideration many of the factors in text that influence the reading process, they are widely used both in practice and in research. Klare (1984) credited Thorndike's The Teacher's Word Book (1921) with being the first milestone in readability as we know it today.

Thorndike's tabulations of the frequency with which words occur in print provided the first way of estimating word difficulty. Lively and Pressey (1923) developed the first of many so-called readability formulas. Many of these formulas are based very simply on percentage of long words and on sentence length (Flesch, 1948; Fry, 1968). Others actually take into consideration the familiarity or difficulty of the words (Dale & Chall, 1948; Spache, 1953).

More sophisticated methods have tried to assess such factors as text structure (Amiran & Jones, 1982). The advent of computer readability software has simplified the use of readability formulas, but the simplest remain the most practical. Of these, The Fry readability formula, with its simple average sentence length and average word length format plus a graph for easy interpretation is "one of the most, if not the most, widely used of all current methods" (Klare, 1984, p. 690).

Determining Reader Ability

The determination of a person's reading ability is commonly established in one of three ways: through performance with graded materials on formal tests, through an informal reading inventory, or through the use of a cloze procedure (Taylor, 1953). Formal tests typically report reading levels in terms of grade

equivalent scores, standard scores, or percentiles. Informal reading inventories and cloze tests allow the identification of three reading levels for each reader.

Individual reading levels. The concept of reading levels for individual readers was developed by Betts (1936; 1954). His original terms for the reading levels were basal, instructional, frustration, and capacity (Betts, 1936). However, the levels later developed by Betts (1954) and still in use today are the independent, the instructional, and the frustration levels.

The independent level is a person's comfortable, easy level of reading. Readers are able to read materials at their independent levels without teacher support. Betts' (1954) criteria for this level are 90% comprehension and 99% word recognition. On a cloze test, students should be able to achieve above 60% exact word replacement in passages at their independent level (Taylor, 1953).

The instructional level is an appropriate level for an instructional situation. Criteria for the instructional level are 75% comprehension and 95% word recognition (Betts, 1954). On a cloze test, the criterion for the instructional level is 40% to 60% exact word replacement (Taylor, 1953).

Material at the frustration level is too difficult for students and should be avoided. The frustration

level is indicated by comprehension of less than 50% and word recognition of less than 90% (Betts, 1954). On the cloze, the criterion for the frustration level is less than 40% exact word replacement (Taylor, 1953).

Establishing reading levels for students and trying to match them with appropriate materials is a widely accepted practice in the teaching of reading. Durkin (1989) recommended establishing individual reading levels in order to choose appropriate textbooks for instruction. Swaby (1989) suggested using a teacher-constructed informal reading inventory to establish reading levels and to supply other information about the students' reading abilities. Taylor, Harris, and Pearson (1988) stated, "For optimal reading growth, it is important that all reading material be at an appropriate level of difficulty" (p. 116). They proceeded to describe the three reading levels, independent, instructional, and frustration in much the same way as they were described by Betts (1954). Gillet and Temple (1990) described the three reading levels, but they included a warning that to believe that a reader has only one single reading level is an oversimplification. Obviously, Betts' concept of reading levels has permeated the entire field of reading. Hoffman et al. (1984) reported, "The notion of appropriate placement in practice materials

has been a part of reading lore for a long time" (p. 381).

Betts' (1936, 1954) original descriptions and criteria are still accepted generally as he proposed them. However, the reading level concept has been further studied by other educators, in particular to discover what exactly is different in the oral and silent reading of students reading at their three different reading levels, what kinds of errors they make, and what the differences are in both quantity and quality of comprehension and strategy use.

Quality of reading at various reading levels. It has been found that as readers move from materials at their instructional level into materials at frustration level, reading either orally or silently, they make different types of errors and there is a tendency to use fewer strategies, such as using meaning and using context clues (Pikulski & Shanahan, 1982). When good readers are placed in materials that are too difficult for them, they display many of the characteristics we associate with poor readers. Their self-correction rate goes down, they have more omission miscues, and they pay less attention to miscues that disrupt meaning or grammatical correctness. The difficulty of the materials keeps them from using active, comprehension-seeking behaviors (Bristow, 1985).

Williamson and Young (1974) found support for these conclusions. They studied the reading performance of intermediate-grade students in instructional- and frustration-level materials. Their conclusions were that students of all abilities adhere more closely to sound and graphic cues than to meaning clues when they are reading at frustration level. They also discovered that students reading at instructional level attended more to paragraph or whole text grammatical and semantic cues, whereas students' use of these cues when reading at frustration level was limited to the sentence or phrase.

D'Angelo and Mahlios (1983) also analyzed the miscues of readers of varying abilities who were reading at their instructional and frustration levels. Their subjects were fifth graders who were classified as good or poor readers on the basis of the comprehension and vocabulary subtests of a standardized test. The researchers found no significant differences between reading miscues based on reading ability when there were appropriate reader-text matches. There were similar insertion and omission miscue patterns for both groups of readers when they were reading under the same reader-text match conditions; that is, when good readers were reading at their instructional or frustration level and poor readers were reading at their respective insructional or frustrational levels.

Leslie and Osol (1978) investigated the oral reading of a heterogeneous group of eighth graders on passages of four different reading levels: 6th-, 8th-, 11th-, and 13th-grade. Their results revealed that when children are reading with at least 95% accuracy they try to use graphic cues as a decoding aid, but that when even superior readers are reading at lower accuracies, these strategies break down. This 95% corresponds to the commonly accepted word recognition criterion for the instructional reading level. Based on the results of their study, Leslie and Osol recommended that students be instructed in materials that they can read with at least 95% accuracy.

Kibby (1979) was working only with disabled readers in his study of oral reading strategies under different reader-text match conditions. He found, of course, a difference in the number of errors, but he also found a different pattern of errors. He concluded that most of the readers in his sample realized that reading is supposed to be meaningful although they were all disabled readers, a group that has been singled out by many researchers as reading on a word-level, decoding basis only (Bristow, 1985; Garner & Kraus, 1981-1982). Kibby found that almost three-fourths of the readers would have been mistakenly identified as readers who do not attend

to meaning if the evaluations had been made on the basis of the more difficult passages.

First graders were subjects in Biemiller's (1979) study of good and poor readers' use of graphic and contextual information. Using passages of increasing difficulty, he found "no clear support for the view that able readers make more use of contextual information than poorer readers" (p. 316).

Studying the effect of the difficulty of material on silent reading is obviously more difficult than the study of its effect on oral reading. However, such methods as analysis of think-aloud protocols (Olshavsky, 1978) and student interviews (Garner & Kraus, 1981-1982) have been used to attempt such an assessment.

Strategy usage in materials of varying difficulties by eleventh graders who were good and poor readers was the focus of a study by Olshavsky (1978). She used analysis of think-aloud protocols to identify 11 reading strategies used by the students. Among the strategies were inference, hypothesis, rereading, and stated failure to understand. She found that both good and poor readers used the strategies less frequently as the difficulty of the materials increased. She hypothesized that all readers judged the difficulty of the material early in their reading and just gave up on the more difficult passages.

In a more recent study, Farris and Triplett (1986) selected fifth graders who read at the fifth-grade instructional level and the fourth-grade independent level. The results indicated that when children read at their independent level, they use more critical thinking skills than when they read at their instructional level. Consequently, the researchers recommended that students be instructed in critical thinking skills in materials at their independent level until they are proficient in the use of such strategies.

Methods of determining reading levels. A primary task in matching readers and texts is the determination of the reading levels of the subjects. In the good reader/poor reader research, students' reading levels have typically been established by using standardized test scores (August et al., 1984; Stanovich et al., 1986; Winograd & Johnston, 1982) or using teacher judgment (Davey, 1988a, 1988b; Garner, 1980).

Two other commonly-accepted methods of assessing individual reading levels are the informal reading inventory and cloze testing. In the informal reading inventory, individual students read short passages and then either give a retelling or answer questions or a combination of both. The scoring procedure is somewhat subjective because there are few absolutely "correct" answers. In addition, there are gaps in the criteria for

establishing the individual levels, and even these criteria have been called into question. In a cloze procedure, words are deleted from a text and students are asked to fill in the words. Only an exact match is accepted, and there are strict criteria for establishing reading levels.

Neither of these two methods of establishing individual reading levels has been used in any of the studies reviewed here. It is understandable that the informal reading inventory has not been used for this purpose because it is an extremely time-consuming procedure.

On the other hand, the cloze procedure is more efficient. The only acceptable answer in a cloze test is the exact match for the word that occurs in the original text, and the scoring criteria are exact. In addition, cloze testing can be done in a group situation.

The cloze procedure is based on the Ebbinghaus Completion Method invented in 1897 by Ebbinghaus (Buros, 1978). Taylor (1953) applied the technique to measuring the readability or difficulty of materials and coined the term cloze. The name derives from the word closure, a term applied by gestalt psychology to the tendency to complete a pattern (Taylor, 1953). In a cloze procedure, words are deleted from a passage according to some

predetermined pattern. The task of the subject is to fill in the blanks.

The cloze procedure is used for diagnosis, for instruction, and for placement. In diagnosis, it is generally used to assess use of context (Gillet & Temple, 1990; Pikulski & Tobin, 1982), but it has sometimes been used to assess overall reading comprehension (Durkin, 1989; Leslie & Osol, 1978). As an instructional strategy, cloze is commonly used to improve use of context (Bridge & Winograd, 1982; Carr, Dewitz, & Patberg, 1989; Jones & Pikulski, 1974; Swaby, 1989). Finally, the cloze procedure is used in placement either to judge the "fit" of a text to a particular reader (Bormuth, 1975; Duffelmeyer, 1983) or to establish individual independent, instructional, and frustration reading levels (Durkin, 1989; Gillet & Temple, 1990; Pikulski & Tobin, 1982).

It is the use of the cloze procedure as a placement device that is of interest here. If an appropriate reader text match is to be made, the independent, instructional, and frustration reading levels of the individual readers must be established. Although the informal reading inventory is probably the most common way of establishing these levels, the cloze procedure has been shown to be correlated with the informal reading inventory (Bormuth, 1968; Rankin & Culhane, 1969) and the

cloze procedure has the advantage of being more efficient to use and more objective in its scoring.

It is well-established that the match between the ability of the reader and the difficulty of the material has an impact on the quality of reading. The cloze procedure and other procedures, such as informal reading inventories, allow us to establish students' individual reading levels so that this match can be made. However, the good reader/poor reader researchers have not attempted to make this match between the text and readers of the same age. In comprehension monitoring studies as well as others, students were given identical materials regardless of their reading ability. The result of this practice is that poorer readers were tested in materials that were relatively more difficult for them than for the better readers.

Purpose of the Study

Therefore, the purpose of this study was to investigate the effect of text difficulty as determined by a readability formula on the comprehension monitoring performance of readers of varying ability. In order to investigate this effect, the comprehension monitoring performance of above-average readers and below-average readers was compared using a typical practice used in the good reader/poor reader research of presenting the same text to all readers, regardless of ability. The results

of this comparison were then compared with the performance of above- and below-average readers placed in different texts at two other difficulty levels, one of which was appropriate, and one of which was difficult, for each group of readers.

Research Questions

The questions addressed in this study were:

1. Is there an interaction effect between material difficulty and reader ability on comprehension monitoring performance?

2. Is there a statistically significant difference between the comprehension monitoring performance of above-average readers and below-average readers overall?

3. Is there a statistically significant difference in reader's comprehension monitoring performance among the three levels of material difficulty?

CHAPTER II

METHODS AND PROCEDURES

The purpose of this study was to determine the effect of text difficulty on the comprehension monitoring of above-average and below-average readers. This chapter contains information about (a) the design of the study, (b) the subjects, (c) the materials, (d) the procedures used in the study, and (e) reduction and analyses of the data.

Design of the Study

The study was conducted using a 2 x 3 factorial design represented in Figure 1. The between-subjects variable was reader ability. There were two levels of this variable: (a) below-average readers, who scored between the 20th and 35th percentiles on the reading subtest of the ITBS and (b) above-average readers, who scored between the 65th and 80th percentiles on the same test. The within-subjects variable was text difficulty. There were three levels of this factor: (a) standard, which placed all the readers in identical second-grade materials that were slightly below the reading level of the below-average readers and considerably below the reading level of the above-average readers; (b) appropriate, which matched readers with different

Reader Ability	Text Difficulty: Standard	Appropriate	Difficult
Below Average	n = 18	n = 18	n = 18
Above Average	n = 18	n = 18	n = 18

Figure 1. Design of the Study. (N = 36)

materials depending on their individual reading levels, third-grade for the below-average readers and fifth-grade for the above-average readers; and (c) difficult, which matched readers with materials above their reading levels, fourth-grade for the below-average readers and sixth-grade for the above-average readers. The interaction between reader ability and text difficulty is referred to in this paper as reader-text match. The dependent variable was comprehension monitoring, which was measured through performance on an error detection task.

Subjects

A pilot test was conducted to assure that the cloze testing procedure would satisfactorily identify the subjects for the study. The materials used were cloze passages adapted from graded materials from Thinking About Science: Focus on Content Reading (Dyer & Lowery, 1988) (see sample in Appendix A).

It was expected that each student's scores would be best on the easiest passage, that is, the passage with the lowest readability, and that their scores would get poorer and poorer as passage difficulty increased. Instead, the scores showed no discernable pattern. Students who did well on the cloze tended to score at instructional level (40% to 60% exact word replacement) on all or most of the passages, while students who scored lower tended to score at frustration level (below 40% replacement) on all or most of the passages. Therefore, it was determined that the cloze testing procedures would not provide the information about reader ability needed for the study.

Based on the results of the pilot testing of the cloze procedure, it was decided to follow the practice of many earlier comprehension monitoring studies and use standardized test scores to select above- and below-average readers for the study. The standardized test scores used were from the reading subtest of the Iowa Tests of Basic Skills (ITBS) (Hieronymous, Hoover, & Lindquist, 1986), which were available in the school files.

In order to assure a separation between the two groups of readers and to avoid using students for whom the materials were too easy or too difficult, both the middle percentiles and the two extremes were avoided.

Olshavsky (1978) used a similar design, choosing as good readers those who scored at stanines 8-9 and poor readers those at stanines 3-4.

The subjects finally selected for the current study consisted of 36 fourth graders from four elementary schools in a midwestern public school district. Students selected as above-average readers were 18 fourth graders who scored between the 65th and 80th percentiles on the ITBS reading subtest, with a mean percentile of 78.06. Their mean grade equivalent score of 5.0 placed them approximately one grade level above their placement in school. The 18 fourth-grade students selected as below-average readers scored between the 20th and 35th percentiles. Their mean percentile was 29.06 and their grade equivalent score was 3.4, approximately one grade level below their school placement. Each group consisted of 8 boys and 10 girls. The study was conducted during early March. Scores used were from the previous fall's testing. The parent permission letter, which was required to be signed and returned, is found in Appendix B.

Materials

The materials used in the error detection task (see Appendix C) were expository science passages adapted from Thinking about Science: Focus on Content Reading (Dyer &

Lowery, 1988). The Fry readability formula was used to determine their readability.

The two standard level passages, "Gold in the Ground" and "An Answer to the Corn Puzzle," had a readability of second-grade. These passages were administered to all students.

The appropriate level passages for the below-average readers, "Life From Soil" and "The Life of a Ladybird," had a third-grade readability. The appropriate level passages for the above-average readers, "Flying Hot Air Balloons" and "The Power of Hurricanes," were fifth-grade passages.

The difficult level passages for the below-average readers, "All About Sharks" and "Getting To Know Snakes," had a fourth-grade readability. The sixth-grade level passages, "Alligator Facts" and "Rainbows Across the Sky," were designated difficult for the above-average readers.

The passages were altered so that five paragraphs in each passage contained a deliberate error, for a total of 10 errors on each level. The errors used were blatant inconsistencies within the text itself. This magnitude of error was chosen because previous research findings have shown that many times not even better readers perform well on comprehension monitoring tasks unless the

errors are blatant. An example of a paragraph containing such a blatant error is given below:

> Water also breaks down rocks when it moves. It may go over waterfalls, through streams or over the elephant in waves. The faster the water moves, the more it will make rocks break into smaller pieces.

The following instructions followed each paragraph:

> Is there a problem with this paragraph?
>
> Yes_____ No_____.
>
> If you answered "Yes," go back and underline the word that seems to cause the problem.

Before the actual data collection began, a pilot test of the error detection materials was conducted to determine whether there was any problem with any particular passage. Fourth-grade students from two intact classrooms in a school not used in the actual study were instructed in the error detection procedure, following the script in Appendix D, and a sample passage (see Appendix E) was done as a group. Their families had received letters requesting them to return and sign them only if they objected to their child's participation in the study (see Appendix F).

After the students had finished the passages, they were interviewed about whether they had found problems with any particular passage and whether there was

anything about the procedure itself that they found confusing. No problems were identified.

Procedures

Before the error detection process began, the six possible orders in which the materials could be administered were identified. The passages could have been administered in these six orders:

1. Standard, Appropriate, Difficult
2. Standard, Difficult, Appropriate
3. Appropriate, Standard, Difficult
4. Appropriate, Difficult, Standard
5. Difficult, Standard, Appropriate
6. Difficult, Appropriate, Standard

Three of those six orders were randomly selected for use in the study. They were (4) Appropriate, Difficult, Standard; (5) Difficult, Standard, Appropriate; and (2) Standard, Difficult, Appropriate. Six above-average readers and six below-average readers were randomly assigned to each of the three orders. This procedure was used to control for any possible effect the order of administration of materials might have on student performance.

The actual study was conducted outside the regular classroom by the researcher. All subjects from a particular school, both above- and below-average readers,

were tested as a group. Testing groups ranged from 11 to 15 subjects.

First, subjects were trained in the error detection task so that it was clear to them. They were shown a sample passage (see Appendix F) and the type of error they were looking for was explained to them. As a group, they orally practiced answering the yes/no question and finding errors and marking them. Then, they were given their individual packets of materials, which had been made up ahead of time. Each packet contained the same two standard reading level passages, which had a readability of second-grade. Each packet also contained two appropriate reading level passages, third-grade readability for the below-average readers and fifth-grade readability for the above-average readers. Finally, the packets contained two difficult reading level passages, fourth-grade readability for the below-average readers and sixth-grade readability for the above-average readers. Students were instructed to do their passages in the order in which they were arranged in the packet, which was the order randomly selected for that student.

In the standard reading level condition, subjects were given two second-grade passages that contained errors as described above. Second-grade passages were used in this study, because this level was below the reading level of all the subjects.

In the appropriate reading level condition, each student was given two passages on his or her appropriate reading level. That is, the below-average readers received a third-grade passage, and the above-average readers received a fifth-grade passage. The levels of the materials corresponded well with the subjects' ITBS grade equivalent scores of 3.4 and 5.0 respectively.

In the difficult reading level condition, each student was given two passages designed to be above his or her appropriate reading level. That is, the below-average readers received a fourth-grade passage, and the above-average readers received a sixth-grade passage.

Data Reduction and Analyses

Scoring was based on ability to detect and underline inconsistencies in the passages. Scores were percentage correct. Correct responses consisted of: (a) answering "Yes" to the question and underlining errors in paragraphs where there are errors, and (b) answering "No" to the question and not underlining anything in paragraphs where no errors were present.

The error detection scores were subjected to a 2 x 3 Analysis of Variance (ANOVA) (Reader Ability x Text Difficulty) with repeated measures on the text difficulty factor. The main effects and all interactions involving the order factor were examined. The two-way interaction

between reader ability and text difficulty (reader-text match) was tested against the null hypothesis of no interaction effect. If a significant interaction was detected, appropriate post hoc comparisons were to be conducted to determine the source of the interaction.

CHAPTER III

RESULTS

The purpose of this study was to determine whether readers' comprehension monitoring performance is a function of the match between reader ability and the difficulty of the text as determined by the Fry readability formula. The interaction between these two factors is conceptualized in this paper as reader-text match. This chapter presents the findings from the analyses of the data.

The data were subjected to a 2 x 3 (Reader Ability x Text Difficulty) Analysis of Variance (ANOVA) with repeated measures on the text difficulty variable. The main effects and all interactions involving the order were examined. The two-way interaction (reader-text match) between text difficulty and reader ability was tested against the null hypothesis of no interaction effect. The between-subjects variable was reader ability as measured by the ITBS reading subtest. The within-subjects variable was text difficulty as determined by the Fry readability formula. Means and standard deviations are shown in Table 1.

Table 1

Comprehension Monitoring Scores for Readers of Two Abilities Under Three Levels of Text Difficulty

	Text Difficulty			
Reader Ability	Standard	Appropriate	Difficult	*M*
Above-Average Readers	.91(.20)	.90(.14)	.84(.17)	.88(.17)
Below-Average Readers	.78(.18)	.77(.18)	.79(.18)	.78(.18)
M	.85(.19)	.84(.16)	.82(.17)	

Note: Numbers given are mean percentages with standard deviations shown in parentheses.

Reader-Text Match

It was hypothesized that text difficulty and reader ability would have an interactive effect on the performance of the subjects. The predicted source of the interaction was that above-average readers would perform better than below-average readers on the second-grade reading level passage only, and that on their appropriate level passages or their difficult level passages, there would be no significant difference because those passages

were of the same relative difficulty for all readers. The predicted interaction was not detected ($F(2, 68)<1.0$). Apparently, there is no differential effect for text difficulty for different levels of reading ability.

Main Effect of Reader Ability

It was hypothesized that the above-average readers would out-perform the below-average readers overall in comprehension monitoring. As would be expected, this hypothesis was upheld. The main effect of reader ability, $F(1, 34) = 5.69$, $p<.05$, was significant. This supported the hypothesis that above-average readers would perform better than below-average readers overall. The above-average readers averaged 88% across all three levels of material and the below-average readers averaged 78%. Standard deviations of .17 for the above-average readers and .18 for the below-average readers indicated a nearly identical degree of variability in the performances of the subjects within each of the two groups. Apparently, it cannot be assumed that as reading ability improves or decreases, variability within groups of readers changes.

Main Effect of Text Difficulty

It was hypothesized that there would be a statistically significant difference in reader comprehension monitoring performance among the three

levels of material difficulty and that readers would perform best on the standard reading level passages. These passages had a readability of second-grade, a level below that of all of the readers in the study. Scores on the appropriate reading level passages, which matched students with passages appropriate for their reading ability, were predicted to be lower because these passages were more challenging. Scores on the difficult reading level passages were expected to be the lowest of the three. Results did not uphold this expectation ($F(2, 68)<1.0$). Indeed, the reading level of the material appeared to have no systematic effect on the performance of the students. When the scores for the two groups of readers were combined, the means for standard, appropriate, and difficult passages were similar: 85%, 84%, and 82%, respectively. The mean scores of the below-average reader group were quite consistent across the sets of materials (see Table 1) on the standard, appropriate, and difficult levels respectively.

Supplemental Analyses

The mean scores alone, however, do not reveal some important within-group characteristics of the data. First, the standard deviations of the two groups of readers were nearly identical, .17 for the above-average readers and .18 for the below-average readers, indicating that the above-average readers were as different from one

another as the below-average readers were from each other.

The distribution of all the scores about the Grand Mean is also somewhat unusual. It would have been expected, given differences in reading ability, that the above-average readers' scores would cluster at the high end of the distribution and the below-average readers' scores would cluster toward the low end. This did not prove to be the case (see Figures 2, 3, and 4). Individual subjects' scores on the individual passages are given in Appendix G. Below-average readers' scores on the standard passages ranged from a low of 17% on the standard level "Gold" passage for subject number 20 to a high of 100% for several subjects on each of the three material levels of difficulty. Scores for the above-average readers ranged from 0% correct for subject number three on the "Corn" passage to a high of 100% for many above-average reader subjects at each of the three levels. The frequency distributions in Figures 2, 3, and 4 show that there was little difference in the distribution of scores between below-average and above-average readers on the task, even on the standard passage. In general, only the clustering of above-average reader scores in the 81-90 and 91-100 ranges differentiates the two groups.

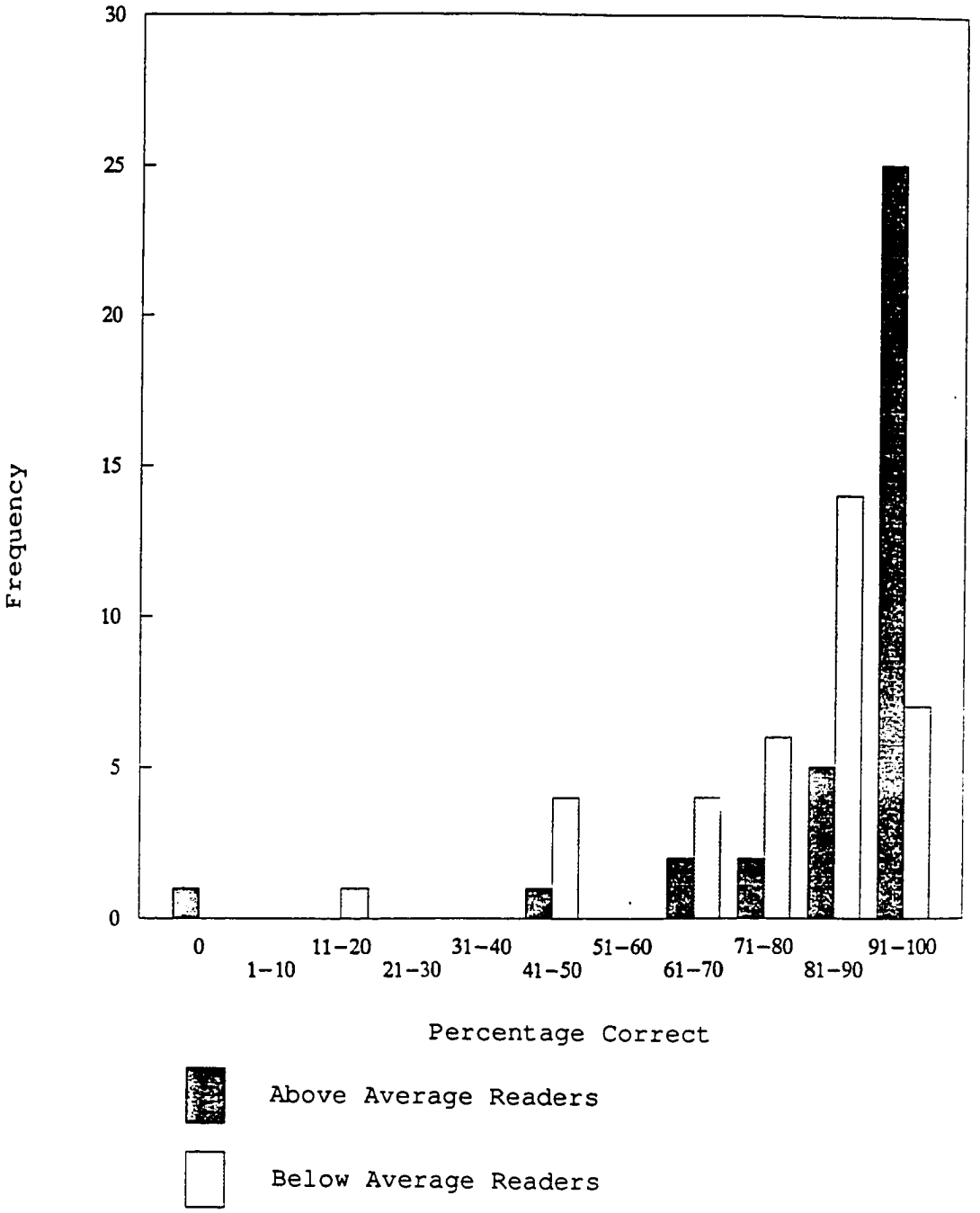

Figure 2. Comprehension Monitoring--Standard Text

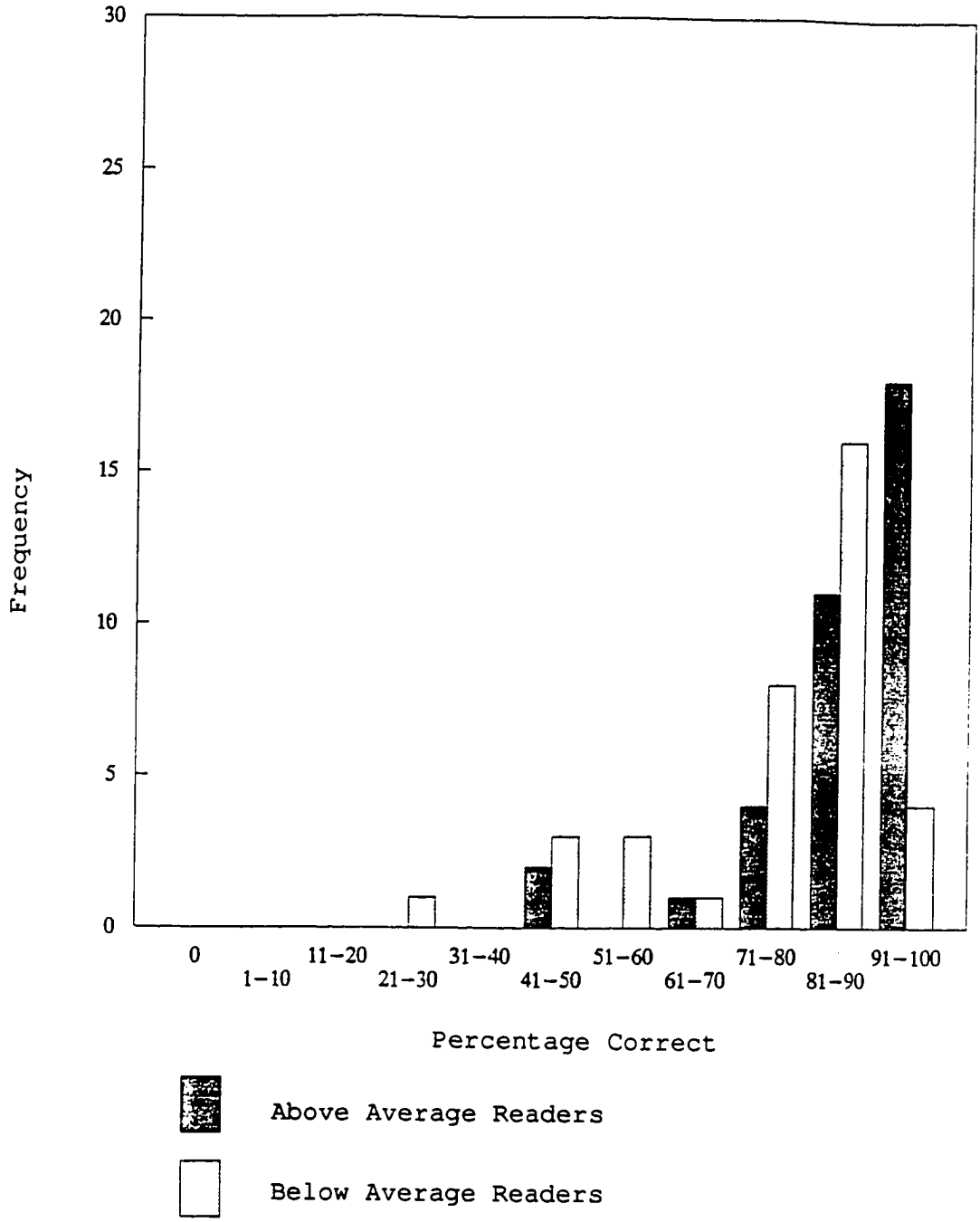

Figure 3. Comprehension Monitoring--Appropriate Text.

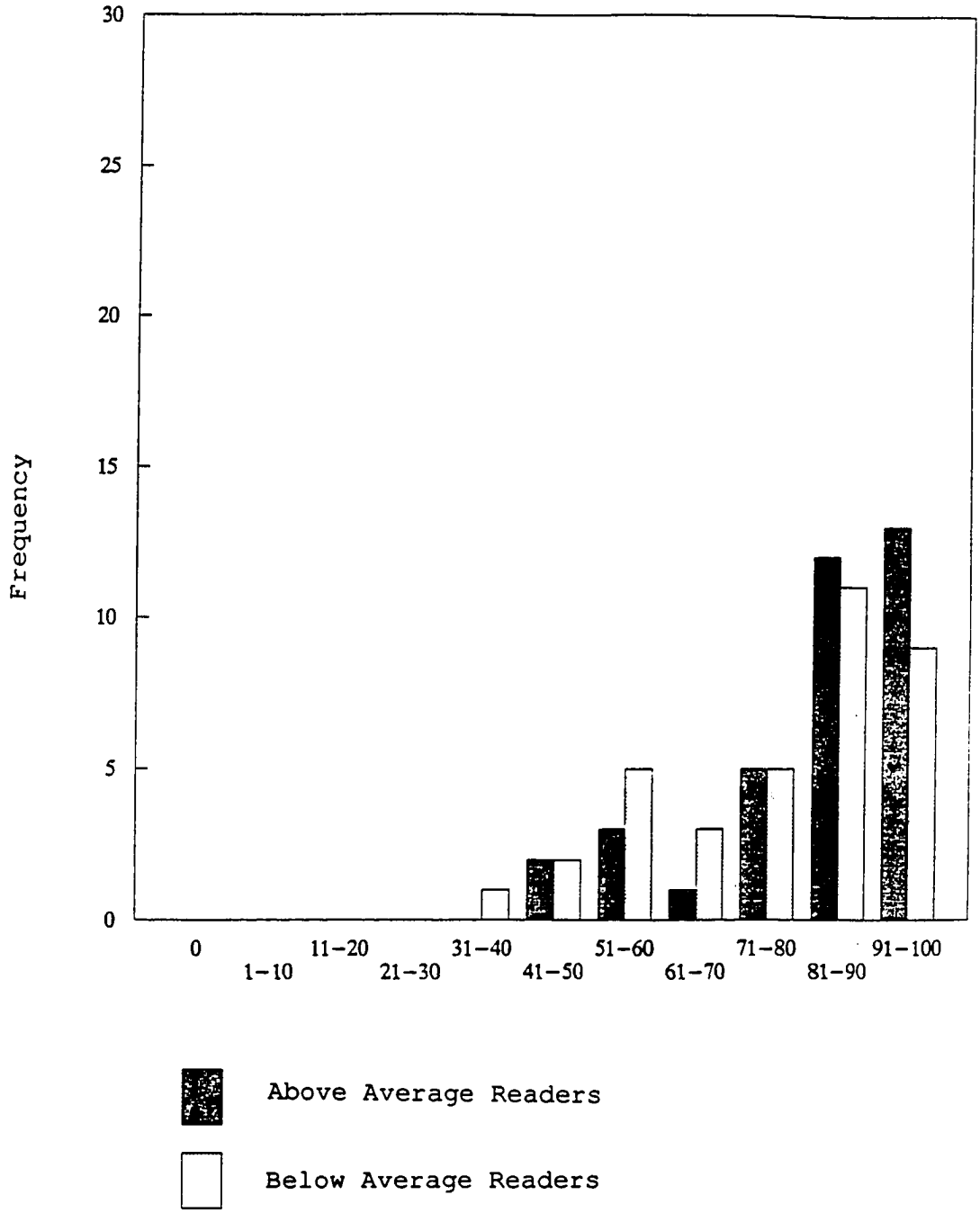

Figure 4. Comprehension Monitoring--Difficult Text.

In most other categories there is a similar representation of both above- and below-average readers.

Even the striking difference in mean performance between the two groups may be reexamined. In an attempt to fully understand the source of the scores, the differences between the groups were examined more closely as a supplement to the basic analysis.

The criteria for a correct answer as described above were very stringent, requiring total accuracy. The only two correct responses were (a) Answering "yes" when there was an error and underlining the error and (b) Answering "no" when there was no error and underlining nothing. However, it is possible that some subjects could have been aware that there was an error; in other words, they could have been able to carry out the first stage of comprehension monitoring, evaluation, and been aware that there was a breakdown in their comprehension, but might have been unable to locate the source of that breakdown.

With this possibility in mind, the errors where a subject answered "yes" when there was an error but underlined the wrong word were examined more closely. It was found that the below-average readers made 43 errors of this kind, accounting for 24% of their total errors while the above-average readers made 14 such errors, accounting for 14% of their total errors. These results supported the possibility that there may be a

developmental level of comprehension monitoring that has not previously been considered. At this level, if it exists, readers may be aware that something is wrong but may be unable to locate the source of the difficulty. This is especially supported by the fact that so many of the below-average readers' errors fell into this category.

To investigate the possibility that students might just be making random guesses, the incorrectly underlined words in the set of errors where students answered "yes" when there was an error but underlined the wrong word were examined to see if they were reasonable (but wrong) guesses, or just wild guesses. The incorrectly underlined words did, for the most part, appear to have some possible rationale for selection. As further evidence that the students were not simply making random guesses, the same incorrectly underlined words were often underlined by more than one student. For example, in one of the standard reading level passages, "An Answer to the Corn Puzzle," the word sand in the following sentence was incorrectly underlined as an error by two above-average readers and two below-average readers: "Sometimes they popped it by putting it into clay pots of hot sand." It is possible that all of these students found that information to be contrary to what they believed could be possible. Another example of such an error is the

mistaken underlining of the word submarine in the following excerpt from "Alligator Facts," a sixth-grade level difficult reading level passage read by the above-average readers: "Sometimes hunters would even snatch alligators from their submarine dens in order to kill them." It is probable that the four above-average readers who underlined the word were unaware that submarine may simply mean underwater.

If these errors, in which a student may be inferred to be aware of a breakdown without being able to locate its source, are added to the number correct of all the students, the mean scores of the two groups of readers across all levels of materials become 83% for the below-average readers and 89% for the above-average readers, considerably closer than the 78% and 88% respectively in the original analysis. The comprehension monitoring performance of the below-average readers may, in fact, be closer to the level of the above-average readers than the scoring system used in the analysis indicated, a possible indication that general reading ability is not necessarily a good predictor of comprehension monitoring ability.

The measure of performance for each level of text difficulty was the sum of scores from two passages that had been equated by the Fry (1968) readability method. In order to assess the success of that method of

selection, and to get a better understanding of the performance of the students, scores on the two comparable passages were examined separately. Scores for the individual passages appear in Appendix G.

There were greater differences among the above-average readers' scores than among the below-average readers' scores. For the above-average reader, individual passage mean scores were 87% and 94% for "Corn" and "Gold," respectively. On their appropriate level passages, the above-average readers' scores were 87% for "Flying Hot Air Balloons" and 93% for "The Power of Hurricanes." A greater difference in scores on two passages of the same relative difficulty was seen on the above-average readers' difficult level passages. The score for "Alligator Facts," a sixth-grade passage that had been judged to be difficult for these readers, was only 77%, as opposed to 91% for the "Rainbows Across the Sky" passage, which was also at sixth-grade level.

Individual passage mean scores for the below-average readers were 80% for "An Answer to the Corn Puzzle" and 76% for "Gold in the Ground," the two standard level passages. On the appropriate level, their scores were 76% and 78% for "Getting to Know Snakes" and "Life From Soil," which were the appropriate level passages. On the difficult level passages, their scores were 80% for "The

Life of the Ladybird" and 77% for "All About Sharks." Clearly, these large differences between passage scores indicated that some factor or factors in the passages other than readability had an impact on the performance of the readers. It also appeared that whatever that factor was or those factors were, the above-average readers were more affected than the below-average readers.

Summary

This study yielded some surprising results. Even the seemingly safe prediction that comprehension monitoring performance would increase as general reading ability increased cannot be reported without discussion. As expected, the above-average readers did, indeed, outperform the below-average readers on the error detection task that was used to assess comprehension monitoring, but a closer look at the data indicated that if a different standard had been applied, the two groups would be closer in performance than it appeared from the analyses used in this study. It appears that some students seemed to be aware that there was an error in a paragraph but were unable to detect its source. This raises the question of whether there may be a developmental stage in comprehension monitoring development where readers may be able to evaluate their comprehension to the extent that they may be aware that

an error exists, but where they may not be able to locate the source of the error.

In addition, reading ability as determined by a standardized test was not a good predictor of comprehension monitoring performance in this study. Although error detection overall was good compared to results from earlier studies, in both groups there were students who did very well and students who did very poorly. The frequency distributions indicated that except for the 81-90 range and the 91-100 range, where more above-average readers were clustered, scores from both groups were generally evenly distributed. Some of the reader factors other than general reading ability that may have influenced performance will be discussed in chapter 4.

The expectation that there would be no significant difference in the comprehension monitoring performance of above-average and below-average readers when relative text difficulty was held constant was not upheld. This is contrary to the indications of earlier studies and brings into question the validity of using readability as measured by a formula to establish text difficulty. Looking at performance on the individual passages that make up each reading level, it appeared that text factors other than simple readability may have had an influence

on performance. Chapter 4 will include a more indepth discussion of some of these factors.

CHAPTER IV

DISCUSSION

This chapter presents a discussion of the findings of this study and offers conclusions about the problem addressed through the study. Implications for theory, suggestions for research, and recommendations for instruction will also be presented.

This study was conducted to determine the effect of text difficulty as determined by the Fry readability formula on the comprehension monitoring performance of above- and below-average fourth-grade readers whose selection was based on their standardized test scores. A handful of earlier studies indicated that the ability to monitor comprehension might be dependent on the match between the text and the reader--that good readers and poor readers alike might be able to monitor their comprehension so long as the material was at an appropriate level for them, but that their performance might diminish as the material became more difficult. Based on that possibility, the subjects in this study were given passages at levels that were matched to their reading levels. The hypothesis that their monitoring performance would get worse as the material got harder was not supported. Indeed, the readability or difficulty

of the material, determined in this study solely by the Fry readability formula, did not seem to have any pattern of influence on the performance of any of the readers. However, something in the passages or within the readers did cause performance to vary from passage to passage, but not in any sequential way. An example is the above-average readers' performance on the two sixth-grade passages that were designated as being difficult for them. On "Rainbows in the Sky," their error detection rate was 91% correct, but on "Alligator Facts," a passage with the same Fry readability, they scored only 71% correct.

Whatever it was that caused reader performance to differ was not predictable to the extent of influencing it in the same direction. On the two standard level passages that were given to all readers, the above-average readers scored better, with a mean score of 94%, on "Gold in the Ground" than they did on "An Answer to the Corn Puzzle," where their mean score was 87%. However, the below-average readers did better on "Corn," 80%, than on "Gold," 76%. A closer examination of any pair of passages on the same readability level shows that individual readers' scores might go up or down from passage to passage, but not according to any discernable pattern. The picture is one of idiosyncratic performance from reader to reader and from passage to passage.

Another hypothesis of this study was that the above-average readers would outperform the below-average readers overall. That seemed a safe assumption. By definition, the better readers should do better on a reading task. They did do better overall, but the pattern of the performance of the individual readers in both groups indicates that the question can neither be asked nor answered so simply. In both groups there were readers who did very well and readers who did very poorly. One would expect when looking at the distribution of the scores to see the better readers clustered toward the higher end of the distribution while the poorer readers clustered toward the lower end. The distributions indicated that this was not the case. Differences appeared primarily in the 81-90 and 91-100 ranges, where there are more above-average readers' scores. Throughout the rest of the distribution, the above-average and below-average readers are fairly equally distributed. It seems that the standardized test scores used to select above- or below-average readers were not a good predictor of comprehension monitoring performance.

As a supplemental analysis, protocols were reexamined to determine the exact nature of students' errors. This reexamination revealed a substantial number of errors which would not have been errors if the

criteria had been different. It appeared that in paragraphs that contained an error, students would often indicate that an error existed by marking "yes," but would then underline as an error a word that was correct but that was, perhaps, contrary to their previous knowledge. As an example, in "An Answer to the Corn Puzzle," two readers from each group underlined the word sand in the following sentence: "Sometimes they popped it by putting into clay pots of hot sand." It may be inferred that students who marked such nonerrors were, indeed, monitoring their comprehension, but that they simply chose what was an error in their minds, rather than the word intended by the researcher as the error. A reanalysis of the data including this group of errors brought the mean scores of the two groups of readers closer together than did the original analysis, which required complete correctness.

In general, the results yielded by this study did not match the expectations. How, then, are these results to be explained? First, the two basic concepts involved in the study must be examined. They are the reading ability of the subjects and the difficulty of the texts.

Reader Ability

The reading ability of the subjects was assessed through a standardized test. Although standardized tests scores are generally considered to be unreliable

indicators of individual student performance (Gillet & Temple, 1990), they could reasonably be expected to correlate to some degree with performance on other reading-related tasks. That did not appear to be the case in this study. Whatever is measured in the ITBS reading subtest did not appear to be very strongly related to the measure of comprehension monitoring used in this study. In other words, the ITBS score was not a good predictor of student performance on the error detection task. Yet, a common assumption, based on earlier studies and on "common sense," is that most poorer readers need instruction in comprehension monitoring and that most better readers do not.

Text Difficulty

The second basic concept in this study was text difficulty. In this study, as in many earlier studies, the assessment of text difficulty was based on a simple, commonly used readability formula. The Fry (1968) readability formula is calculated solely on word length and sentence length. Earlier studies have shown that varying levels of difficulty of material based on this type of readability influenced reading performance. Why then did it appear to have no pattern of influence in this study?

In order to attempt to answer this question, it is necessary to look at some of the criticisms of

readability formulas. When readability formulas were developed, they tended to be seen as a panacea for all illnesses in reading. That enchantment persisted for many years, but recent emphasis on the nature of text comprehension as an interaction between the text and the reader has left open to question the assumptions on which readability formulas are based. The assumption that readability is something that resides solely in the text is no longer accepted without question (Gillet & Temple, 1990). Cadenhead (1987) stated, "the background, interests, and aspirations of children affect their performance in ways that we have not yet learned to assess" (p. 438). Davison and Kantor (1982) believed that "reading difficulty may be affected by the purposes and background of the reader and the inherent difficulties of the subject matter; it is not just a function of measurable properties like length and vocabulary" (p. 189).

Factors That Affect Reading Performance

There are many factors, both within the reader and within the text, that are known to influence the reading process and that are not taken into consideration by the readability formulas. Reader factors include prior knowledge of the text topic (Garner, 1990; Gillet & Temple, 1990; Guthrie, 1983; Johnston, 1983; Lipson & Wixson, 1986; Richgels, 1982; Wilson & Gambrell, 1988),

interest in the topic (Durkin, 1989; Johnston, 1983), knowledge of text structure (Johnston, 1983; Raphael et al., 1981), and such affective factors as motivation (Klare, 1984; Taylor et al., 1988), self-concept (Lipson & Wixson, 1986), attribution (Garner, 1990; Lipson & Wixson, 1986), attitude (McKenna & Kear, 1990) and risk-taking ability (Gentile & McMillan, 1987; Lipson & Wixson, 1986).

Factors within the text that influence the reading process include topic of text (Singer & Donlan, 1989), text structure (Garner, 1990; Gillet & Temple, 1990), and a multitude of features that make a text considerate or friendly, including organization and coherence (Durkin, 1989; Klare, 1984), concept density (Johnston, 1983; Wilson & Gambrell, 1988), and explication or directness (Durkin, 1989; Johnston, 1983). With so many factors ignored, it is no wonder that readability formulas do not present a complete picture of true readability.

It has been argued that decreasing the difficulty of material by rewriting it so that it has a lower readability may actually make the material more difficult for students. In their analysis of the effects of adaptation of text on its readability, Davison and Kantor (1982) concluded that the assumptions on which readability formulas are based may not always hold true. For example, sentence length does not always contribute

to complexity. Sometimes breaking a complex sentence into parts destroys the relationship between its parts, requiring additional words to be added or, if ignored, leaving the reader without support in understanding that relationship. A difficult vocabulary word may not have an acceptable synonym, so that additional words are required and clarity is lost. In short, using readability formulas to adapt text does not make text easier to read. In fact, the opposite may be the case.

Shortcomings of readability formulas indicate that, although text difficulty as determined by readability formulas appeared to influence student performance in some earlier studies, there are many other more powerful factors that may affect the performance of the readers. It clearly cannot be assumed that these factors are based solely in the text, as is assumed by readability formulas. What is now known about the interactive nature of reading points toward answers that involve the readers, the texts, and the interaction between them.

The findings in this study provided evidence that above-average and below-average readers' comprehension monitoring performance is affected by the text but not by the Fry readability of the text as was hypothesized. These results do not support previous research (Erickson et al., 1985; Latvala, 1989; Paris & Myers, 1981). It is, therefore, appropriate to discuss other factors

within the text or within the readers and interactions between reader and text factors that may have caused the differences in performance evidenced in this study.

Factors Within the Reader

Some of the most important factors affecting reading performance are found within the reader. These factors may be specific to each passage and each reader, perhaps accounting for some of the variability within individual performances in this study, or they may be more generalized, affecting the student's overall performance. Some of the reader factors to be discussed here are those factors that need to be considered in matching specific readers with specific texts. These are prior knowledge and schemata, interest in the topic, and knowledge of the structure of the text. Several affective reader factors such as motivation, attitude, and attribution that may have a generalized effect on performance will be briefly considered, but a lengthy exploration of these factors is not within the scope of this discussion.

Prior knowledge and schemata. Reading comprehension is largely a process of combining the known with the new to compose a meaning (Pearson & Johnson, 1978). The known in this case refers to the reader's schema for a particular topic. This concept is also referred to as prior knowledge (Johnston, 1983; Klare, 1984), background

knowledge (Durkin, 1989), or world knowledge (Gillet & Temple, 1990).

Prior knowledge has been found to have a powerful influence on reading comprehension and related tasks. For example, in their 1983 study of text comprehension, Freebody and Anderson found that familiarity with the topic of the text accounted for almost three times as much variance as did vocabulary difficulty. In Recht and Leslie's (1988) study, the high-ability readers with low knowledge of text topic did no better on a memory task than did low-ability readers with low knowledge of the topic. Even the high-ability readers with high prior knowledge did no better than low-ability readers who also had high topic knowledge. Anderson and Acker's (1984) results are similar. In their study, when there was no significant amount of prior knowledge, good and poor comprehenders performed similarly on comprehension tasks. Afflerbach (1990) found that the prior knowledge of his subjects influenced a range of comprehension processes including comprehension monitoring.

Prior knowledge is one of the factors that is thought to override readability as judged by a formula. When readers have higher prior knowledge they are less affected by increases in readability difficulty (Caldwell, 1985; Klare, 1984), and they have less need for strategy use (Garner, 1990).

Indeed, prior knowledge is so powerful that it can have an adverse effect on comprehension if the prior knowledge activated before reading is incompatible with the information in the text (Alvermann, Smith, & Readence, 1985). In some cases, prior knowledge can override text information so that readers respond to questions with information from their knowledge base even when it is not supported by the text (McKeown, Beck, Sinatra, & Loxterman, 1992).

Prior knowledge is acknowledged to play a powerful part in comprehension. It can even compensate somewhat for less well-constructed texts, giving students a chance of comprehending (Wilson & Anderson, 1986). However, it can not completely compensate for textual inadequacies (McKeown et al., 1992). Based on her findings about the relationship between prior knowledge and text structure, Yochum (1991) concluded that the effect of prior knowledge may vary depending on the task, the information to be learned, and other reader factors such as ability.

Although prior knowledge alone cannot compensate for all deficiencies, educators agree that it plays an important role in comprehension (Durkin, 1989; Gillet & Temple, 1990; Taylor et al., 1988). "All educators know that it is easier for their students to comprehend a passage whose subject is familiar to them. Good teachers provide background information before assigning reading

on unfamiliar subjects" (Richgels, 1982, p. 61). Brown (1980) contended that a skilled reader is an active, thinking participant who seeks to get meaning from text by constructing a new knowledge out of selectively combined text information and prior knowledge.

In the current study, it is possible that the unpredictability in readers' performance may be accounted for in part by their knowledge of the particular topics of the specific passages. Prior knowledge is a within-reader factor that interacts with the text factor of text topic. In other words, it is the match between the reader's prior knowledge and the topic of the text that is crucial.

Interest. Interest is a reader factor that is difficult to separate from prior knowledge (Leslie & Caldwell, 1990). Better comprehension of high interest text may be due to increased knowledge of the topic. In other words, it may be that interest in a topic leads to an effort to acquire knowledge about that topic or that knowledge creates interest. Klare (1984) called it "an interesting chicken-egg conundrum" (p. 725).

Guthrie (1981) found that when prior knowledge and reading ability are accounted for, interest does not significantly influence comprehension. However, Baldwin, Peleg-Bruckner, and McClintock (1985) found that interest

and prior knowledge were not correlated, but they speculated that this relationship may vary, depending on whether the subject is a child or an adult. They believed that adults may be freer to develop knowledge in their interest areas while children would have knowledge in topics in which they have been schooled, regardless of their interest.

Whatever its relationship with prior knowledge, interest in the topic is considered a factor that influences comprehension (Asher, 1980; Durkin, 1989; Lipson & Wixson, 1986; Taylor et al., 1988). The better the match between a student's interests and the topic of the text, the better the student's performance will be.

Although the passages used in this study were all on scientific subjects, there was a wide variety of topics. It is possible that a student might be bored with the passage about corn, a passage with a second-grade Fry readability, but would be fascinated by the one about hurricanes, a sixth-grade Fry readability level passage. The student's interest may have caused him or her to perform better on the "more difficult" passage than on the "easier," but for the student, less-interesting passage, and thus accounted for some of the seemingly patternless variability within individual students' performances.

Knowledge of text structure. Each particular type of text has its own structure. Narrative text follows a story grammar structure with a setting, characters, a problem and solution, and a conclusion. It is this structure that students are most familiar with and, therefore, it is also this structure that is easier for most of them to read (Lipson & Wixson, 1986). Younger students have been found to lack knowledge about expository text structure (Garner & Gillingham, 1987). Even students who succeed in reading in the early grades may begin having difficulty at fourth or fifth grade if they have not been taught the structure of expository text (Richards, 1978). Up until that time, they have probably received most of their reading instruction in narrative texts, but beginning at fourth grade they must read more expository text (McCormick, 1989). Indeed, even many older students lack knowledge of nonfiction or expository text structure, and their comprehension of expository text is adversely affected by that lack of knowledge (Marshall & Glock, 1978-1979).

Singer and Donlan (1989) described expository texts as those that "contain explanations of objects, events, situations or procedures for carrying out activities." Expository selections are generally more difficult to comprehend than narrative because of such factors as concept or proposition density (McCormick, 1989).

Some of the rhetorical structures or organizational patterns common to expository text are cause and effect, comparison and contrast, and question and answer. In good expository text, there are signals that help the reader follow the organization and understand the relationships. These may be classified as semantic signals such as first, second, and finally, or syntactic signals such as topic sentences and summaries (Piccolo, 1987).

Knowledge of text structure may have had a specific effect on individual students' performances in this study as well as a generalized effect. Some individual differences may be explained by the reader's knowledge of the specific rhetorical structure (i.e., cause and effect or comparison and contrast) of a specific passage. In addition, a student's general knowledge of and experience with expository text would affect how that student did overall.

Affective reader factors. There are many affective reader factors that tend to have a generalized effect on reading performance. Those to be briefly discussed here are motivation, attitude, and attribution.

Motivation has a major role in determining a student's performance (Klare, 1984). Wilson and Gambrell (1988) called it a powerful variable in the reading comprehension process. Motivation is kind of "inner

push." The more strongly we want or need something, the more motivated we are. Motivations can be internal such as feelings of satisfaction or competence or external such as material rewards or avoidance of punishment (Taylor et al., 1988). Expressed more simply, motivation is simply how badly someone wants to do something. Winograd and Paris (1989) said, "developing a motivational agenda is crucial to improving reading instruction" (p. 32).

Another affective factor within the reader is attitude or predisposition toward reading, toward the topic, and toward the task. Wixson and Lipson (in press) called the student's attitude toward reading "a central factor affecting reading performance." McKenna and Kear (1990) reported that there is a long history of research in which attitude and achievement have been consistently linked.

Attribution has also been found to play a significant part in student performance. The term attribution refers to explanations people develop for their success or failure (Weiner, 1974). Success or failure may be attributed to such internal factors as ability or effort or to external factors such as task difficulty or luck. A related concept is learned helplessness (Bristow, 1985). This is a phenomenon usually associated with low-achieving students who

believe that their reading failures are attributable to their own lack of ability and that there is, therefore, nothing they can do about them. Higher-achieving students tend to believe that more effort on their part will result in better performance. In this study, attribution, although it certainly may have affected performance, would be a general factor that would affect a student's performance on all of the passages rather than differing from passage to passage.

These affective reader factors, motivation, attitude, and attribution, may quite possibly have influenced the performances of the readers in this study. However, their effect would tend to be an overall, generalized one rather that affecting student performance from passage to passage.

Factors Within the Text

Having considered the reader factors that affect performance on reading and reading-related tasks, we must now look at factors within the text that make it more or less difficult or simply a better or worse match for an individual student. As in the discussion of reader factors, some text factors are specific to each passage while others may be generalized, affecting subjects' performance overall. Idiosyncratic text factors may have contributed to the erratic results from passage to passage seen in this study. Among the text factors that

need to be considered in matching individual texts with individual readers are the topic of the text and the structure of the text. Several others, such as organization, cohesion, explication or directness, and concept density, come under the general heading of text considerateness or friendliness (Singer & Donlan, 1989).

Text topic. The topic of a passage interacts with the prior knowledge (Durkin, 1989) and the interests (Lipson & Wixson, 1986) of the reader. Students in this study who had extensive prior knowledge about or interest in sharks may possibly have performed better on the difficult "All About Sharks" passage given to the below-average readers than on a passage with a lower readability but perhaps less interest such as "Life From Soil."

Text structure. As mentioned above, there are several rhetorical structures that are commonly used in expository text. Some of these are sequence, listing, comparison and contrast, cause and effect, problem and solution, and question and answer (Singer & Donlan, 1989). Text structure is a factor that "matches" with the reader's knowledge of text structure. If a specific text has a rhetorical structure with which the reader is not familiar or not adept, there may be said to be a "mismatch" that will adversely affect reading performance.

The passages in this study have not been analyzed to determine their rhetorical structure. However, the significant question is not, "What are the rhetorical structures of the passages in this study?" but rather, "How close is the fit between these passages and the text structure knowledge of each reader?" This is a difficult question to answer, and it was not accounted for in this study. Certainly this factor may have affected the performance of the subjects.

Text considerateness. Beck and McKeown (1986) described considerate text as being "designed to maximize the possibility for the reader to gain information from text and to establish relations among concepts" (p. 129). Among the characteristics of considerate text are organization, coherence or cohesion, explication or directness, and concept load.

Better organized passages are more logically arranged and place fewer demands on the reader. The reader is required to do less recalling of background information, inferring, perceiving relationships, or drawing conclusions (Singer & Donlan, 1989). Well-organized text aids a student's comprehension and retention (Durkin, 1989). This is particularly true when the reader is aware of text structure and the text is well-organized to follow a rhetorical schema such as cause/effect or comparison/contrast (Sawyer, 1991).

Coherence or cohesion is related to organization and is frequently mentioned as a text characteristic that affects the "friendliness" of the text (Durkin, 1989; Lipson & Wixson, 1986). According to McKeown et al. (1992), "coherence is the extent to which the sequencing of ideas in a text makes sense and the extent to which the language used to present those ideas makes the nature of the ideas and their relationships apparent" (p. 79). The more coherent a text is, the easier it is to understand.

It is possible that the passages used in this study may vary significantly in their organization and coherence. If so, that could account for some of the variability from passage to passage within the performances of the individual readers.

The more directly stated or explicit information is, the easier it is to understand. Less explicit or direct materials require more inference on the part of the reader, requiring him or her to "fill in the blanks" to understand relationships and generalizations (Durkin, 1989; Johnston, 1983). Related to this characteristic is the availability within the text of definitions for terms that may be unfamiliar to the reader (Durkin, 1989; Garner, 1987). Thus, the less explicit or direct text is, the more inferences are required of the reader, and the less considerate or friendly the text becomes.

Content area material such as the science passages used in this study sometimes has a high concept density or load. "Concept density refers to the concentration of concepts within a given unit of text" (Wilson & Gambrell, 1988, p. 88). This makes the material more difficult for students, especially if they lack prior knowledge of the topic.

If the passages used in this study were to be analyzed for concept density, it might be found that they vary on that characteristic. If so, that could account for some of the within-student variability observed.

The passages in this study have not been analyzed for any of the considerate text characteristics. Since they all came from the same series of materials, it is possible, but not necessarily true, that they would be similar in their considerateness. If so, the effect of this characteristic would be generalized. If, however, they vary greatly in their considerateness, that could account for some of the unpredicted results of this study, specifically the variations in performance across passages, even within a readability level.

Explanations for Nondetection

Garner (1987) reported that one-third to two-thirds of all errors were undetected in the studies she reviewed. The error detection rate in this study was considerably higher. Combining all student protocols,

there were 1602 total paragraphs in this study. Of that number, 522 paragraphs contained no errors. Students received credit for indicating that there were no errors by answering "no" to the question following each paragraph. The remaining 1080 paragraphs, 540 for each group of readers, contained one error each. Of these 1080 errors, 855, or nearly 80%, were detected. The below-average readers detected 73% of their errors and the above-average readers detected 86% of theirs. Indeed, if the errors where students were aware of a difficulty but unable to locate its source are included as correct answers, the error detection percentages become even higher, 81% for the below-average readers and 88% for the above-average readers. Nevertheless, it is useful to consider why some errors went undetected.

One possibility is that the students' comprehension was not interrupted by the errors. They may have unconsciously "fixed up" the errors as they read. Since the errors were single words, it would be quite possible to understand the whole paragraph without noticing that one word was out of context.

Another kind of "fix up" may have occurred in those cases where students found what they thought were errors because the information was in conflict with their prior knowledge. The subjects had been told that there could be no more than one error in each paragraph.

Consequently, they may have simply not looked any farther, once they thought they had identified an error.

Another explanation may be that when they encountered passages that were difficult for them their cognitive systems became overloaded. If, indeed, lack of prior knowledge was an important factor in determining the difficulty of a particular passage for a particular student, that cognitive overload could have occurred on those passages for which an individual student had insufficient prior knowledge.

It is also quite possible that the students had not received instruction in comprehension monitoring. To some in both groups comprehension monitoring may come naturally, but the others may need instruction or may not have benefited from the instruction they have received.

A final possibility is that some students may have rushed the task. Even though they were told that the task was not timed, there would naturally be a certain amount of pressure from observing those around them. If they saw that others were finishing, they might have rushed to finish also.

Implications for Theory

The clearest message delivered by the results of this study is that readers must be considered as individuals. The extreme complexity of the reading process and the uniqueness of individual readers as they

interact with individual texts must lead to a theory of reading that emphasizes the individual, not the group. Not only each individual, but each reading act must be considered separately, and all characteristics of the reader, the text and the interactions between them must be considered.

Almost as importantly, we must move away from any "quick-fix" solutions to reading problems that consider a below-average reader to be a jigsaw puzzle with one piece missing. Again, the complexity of the reading process must be emphasized.

Suggestions for Research

This study provides some unexpected insights about the effects of text difficulty on above- and below-average fourth grade readers. It brings the concept of readability as measured by a readability formula into serious question particularly as it pertains to expository text. It also calls into question the use of standardized test scores to assume proficiency in reading-related skills such as comprehension monitoring. Several ideas for future research are indicated:

1. Future comprehension monitoring studies should address the question of reader-text match, not by matching standardized test reader ability with readability formula text readability, but by attempting to account for such factors as the match of text topic

and reader's prior knowledge and/or interest and the match between the rhetorical structure of the text and the reader's knowledge of text structure. The considerateness of the text should also be assessed and taken into account as well as generalized affective reader factors such as motivation and attribution.

2. A comprehension monitoring study designed to take into account the factors mentioned above should be conducted comparing reader performance with narrative text and reader performance with expository text. It may be found that such factors as the reader's prior knowledge have a greater effect on performance with one type of text than on the other.

3. There is still a need to find better ways of assessing reading ability in studies such as this one. Informal reading inventories are a possibility, if the time required to administer them is available. Further investigation of the cloze procedure as a reading level assessment should also be conducted. To assure comparability, perhaps the same passages could be used for both the cloze testing and the error detection task.

4. Comprehension monitoring has previously been thought to have two parts, evaluation (check-up) and regulation (fix-up), with some educators adding a maintenance step. From the results of this study, it appears that some readers at times may be aware of a

difficulty (evaluation) without being able to locate the source of that difficulty. It should be investigated whether this may be a stage in the development of comprehension monitoring ability.

5. Studies such as this one, no matter how carefully they are planned, place subjects in artificial situations. They are aware that they are being studied, which inevitably has an impact on their performance. Future studies should attempt to assess comprehension monitoring and other variables in more naturalistic settings, collecting observational information and talking with readers of all levels of reading proficiency about what they actually do when they read.

Recommendations for Instruction

This study indicated that the general reading ability of a reader does not necessarily indicate his or her comprehension monitoring ability. It is ill-advised to assume that a below-average reader needs instruction in comprehension monitoring, nor should we assume that an above-average reader does not need such instruction. Comprehension monitoring is an important reading comprehension strategy, not just for those who score high on standardized tests, but for all readers.

In fact, in this study, reading ability as measured by a standardized test did not prove to be a good predictor of comprehension monitoring performance.

Several of the below-average readers performed quite well on the error detection task while some of the above-average readers did poorly. These results supported the conclusion of August et al. (1984), "Teachers would be well advised to assume that children who score at or above grade level on standardized comprehension tests still need instruction in comprehension monitoring" (p. 48) and of Pearson and Gallagher (1983) who stated, "training either in strategy use or monitoring is beneficial, often to the lower achievers but more often to all students" (p. 337). Comprehension monitoring instruction should be offered to all students so that they may become independent, strategic readers who reach their full potential.

Indeed, the practice of basing any kind of instructional decisions for individual students on the results of standardized achievement tests must also be questioned. These tests were never intended to be used for these purposes. Instructional decisions must be based on a variety of factors, including everyday classroom performance, collection of work samples, assessments that are congruent with the curriculum of the classroom, teacher observation, and on student self-evaluation.

The current trend toward universal testing must certainly be questioned. The results of this and other studies indicate that each reading act is unique, and that the factors affecting it are extremely complex and interactive. Universal testing, where all students take the same test, simply cannot give a clear picture of reading ability for either individuals or groups.

Although simple readability as determined by a readability formula proved inoperative in this study, the concept of readability need not be completely discarded. Readability formulas give only a rough estimate of text difficulty (Hansell, 1976), but they can be useful for just that purpose (Gillet & Temple, 1990). As a quick screening device, they may be useful to busy teachers, but the many other factors that determine the difficulty of a particular text for a particular student must always be considered.

Summary

This study focused on the match between the reader and the text. The results indicated that this match is much more complicated than the simple matching of reader ability as determined by a standardized test with text readability as determined by a readability formula. Primary areas where reader-text match must be considered are (a) the match of the topic of each specific text with the individual reader's prior knowledge of and/or

interest in that topic and (b) the match of the text structure of a specific text with the individual reader's knowledge of that text structure.

Results of this study indicate the importance of using what we already know about assessing prior knowledge, about building prior knowledge when necessary, and about matching readers with texts for which they have a satisfactory knowledge base. We need to find and use more ways of knowing what readers knew before reading to understand what they comprehend during reading (Steele, 1985). Using an instrument such as the Qualitative Reading Inventory (QRI) (Leslie & Caldwell, 1990) that includes assessment of prior knowledge will assist reading educators in assessing prior knowledge in a diagnostic situation. Similar procedures should be used in the classroom. Prior knowledge may be assessed in several different ways (Holmes & Rosser, 1987). It may be assessed through free association with words or concepts closely related with the topic, through prediction based on the title or topic, through a free recall of all a student knows about a topic, through multiple choice or unstructured questions, or through discussion. Interest inventories and careful observation of and conversation with students can give teachers information about students' interests.

However, assessing prior knowledge and interest is not enough. The results of such assessment must be taken into consideration during instruction either through prereading activities or in the amount of support given during instruction. Finally, students must be taught and encouraged to activate their own prior knowledge before and during reading and to build a knowledge base for themselves when necessary.

Instruction in various text structures is also crucial. A student's knowledge of a particular text structure affects both retention of the information in a passage (Englert & Hiebert, 1984; McGee & Richgels, 1985) and comprehension of that text (Gillet & Temple, 1990; Horowitz, 1985). Teaching students story grammar helps them to understand narrative material, but they must also learn about the rhetorical patterns of expository text. This instruction should begin, not at fourth or fifth grade or in junior high school, but in the lower elementary grades (Olson & Gee, 1991).

Finally, the results of this study offer yet another argument for the viewing of reading as a dynamic, interactive process. The reading process is much too complex for us to look for a single variable to serve as the missing puzzle piece that will fix "incomplete" readers, and it is much too complex for us to assume that a reader who scores high on a standardized test will be a

"complete" reader in all situations or that a reader with a low score will be deficient in all areas.

It is time to abolish the stigma attached to being a (dis)abled reader who is forced to become a member of a distinct social group (Lipson & Wixson, 1986). The complexity of the reading process leads to an understanding of the need for informed instruction and facilitation of learning which equally invites all readers to become "members of the literacy club" (Smith, 1988).

REFERENCES

Afflerbach, P. P. (1990). The influence of prior knowledge on expert readers' main idea construction strategies. Reading Research Quarterly, 25, 31-46.

Allington, R. L., & Strange, M. (1978). Word prediction of good and poor readers. In P. D. Pearson & J. Hansen (Eds.), Reading: Disciplined inquiry in process and practice (pp. 56-60). (Twenty-seventh yearbook). Clemson, SC: The National Reading Conference.

Alvermann, D. E., Smith, L. C., & Readence, J. E. (1985). Prior knowledge activation and the comprehension of compatible and incompatible text. Reading Research Quarterly, 20, 420-436.

Amiran, M. R., & Jones, B. F. (1982). Toward a new definition of readability. Educational Psychologist, 17, 13-30.

Anderson, B. (1975). Cognitive psychology: The study of knowing, learning, and thinking. New York: Academic Press.

Anderson, I. H. (1937). Studies in the eye-movements of good and poor readers. Psychological Monographs, 48, 1-35.

Anderson, O. S., & Acker, R. L. (1984). An investigation of the effects of prior knowledge on comprehension of discourse by good and poor comprehenders. Reading World, 24, 87-95.

Asher, S. R. (1980). Topic interest and children's reading comprehension. In R. J. Spiro, B. C. Bruce, & W. F. Brewer, (Eds.), Theoretical issues in reading comprehension (pp. 525-534). Hillsdale, NJ: Lawrence Erlbaum.

August, D. L., Flavell, J. H., & Clift, R. (1984). Comparison of comprehension monitoring of skilled and less skilled readers. Reading Research Quarterly, 20, 39-53.

Babbs, P. J. (1984). Monitoring cards help improve comprehension. The Reading Teacher, 38, 200-204.

Babbs, P. J., & Moe, A. J. (1983). Metacognition: A key for independent learning from text. The Reading Teacher, 36, 422-426.

Backman, J. E., Mamen, M., & Ferguson, H. B. (1984). Reading level design: Conceptual and methodological issues in reading research. Psychological Bulletin, 96, 560-568.

Baker, L. (1985). How do we know when we don't understand? Standards for evaluating text comprehension. In D. L. Forrest, G. E. MacKinnon, & T. G. Waller (Eds.), Metacognition, cognition and human performance (pp. 155-205). New York: Academic Press.

Baker, L., (1987). Children's effective use of multiple standards for evaluating their comprehension. Journal of Educational Psychology, 76, 588-597.

Baker, L., & Anderson, R. I. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. Reading Research Quarterly, 17, 281-294.

Baker, L., & Brown, A. L. (1984a). Cognitive monitoring in reading. In J. Flood (Ed.), Understanding reading comprehension (pp. 21-44). Newark, DE: International Reading Association.

Baker, L., & Brown, A. L. (1984b). Metacognitive skills in reading. In P. D. Pearson (Ed.), Handbook of reading research (pp. 353-394). New York: Longman.

Baldwin, R. S., Peleg-Bruckner, Z., & McClintock, A. H. (1985). Effects of topic interest and prior knowledge on reading comprehension. Reading Research Quarterly, 20, 497-504.

Beck, I., & McKeown, M. G. (1986). Instructional research in reading: A retrospective. In J. Orasanu (Ed.), Reading comprehension from research to practice (pp. 113-134). Hillsdale, NJ: Lawrence Erlbaum.

Beebe, M. J. (1980). The effect of different types of substitution miscues on reading. Reading Research Quarterly, 15, 324-336.

Betts, E. A. (1936). The prevention and correction of reading difficulties. Evanston, IL: Row, Peterson.

Betts, E. A. (1954). Foundations of reading instruction. New York: American Book.

Biemiller, A. (1979). Changes in the use of graphic and contextual information as a function of passage difficulty and reading achievement level. Journal of Reading Behavior, 11, 307-318.

Bormuth, J. R. (1968). Cloze test readability: Criterion referenced scores. Journal of Educational Measurement, 5, 189-196.

Bormuth, J. R. (1975). The cloze procedure: Literacy in the classroom. In W. D. Page (Ed.), Help for the reading teacher: New directions in Research (pp. 68-72). Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills.

Bransford, J. D., Vye, N. J., & Stein, B. S. (1984). A comparison of successful and less successful learners: Can we enhance comprehension and mastery skills? In J. Flood (Ed.)., Promoting reading comprehension (pp. 68-72). Newark, DE: International Reading Association.

Bridge, C. A., Tierney, R. J., & Cera, M. J. (1978). Inferential operations of children involved in discourse processing. In P. D. Pearson & J. Hansen (Eds.), Reading: Disciplined inquiry in process and practice. (Twenty-seventh yearbook). Clemson, SC: The National Reading Conference.

Bridge, C. A., & Winograd, P. N. (1982). Readers' awareness of cohesive relationships during close comprehension. Journal of Reading Behavior, 14, 299-312.

Briggs, P., & Underwood, G. (1987). The nature of reader ability differences in lexical access. Journal of Research in Reading, 10, 57-74.

Bristow, P. S. (1985). Are poor readers passive readers? Some evidence, possible explanations, and possible solutions. The Reading Teacher, 39, 318-325.

Brown, A. L. (1975). The development of memory. Knowing about knowing and knowing how to know. Advances in Child Development and Behavior, 10, 103-151.

Brown, A. L. (1980). Metacognitive development and reading. In J. Rand, B. B. Spiro, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension (pp. 453-482). Hillsdale, NJ: Lawrence Erlbaum.

Brown A. L., Armbruster, B. B., & Baker, L. (1986). The role of metacognition in reading and studying. In J. Orasanu (Ed.), Reading comprehension: From research to practice (pp. 49-76). Hillsdale, NJ: Erlbaum.

Bryant, P., & Goswami, U. (1986). Strengths and weaknesses of the reading level design: A comment on Backman, Mamen and Ferguson. Psychological Bulletin, 100, 101-103.

Buros, O. R. (1978). Eighth mental measurements yearbook, Vol. 11. Highland Park, NJ: Gryphon Press.

Cadenhead, K. (1987). Reading level: A metaphor that shapes practice. Phi Delta Kappan, 68, 436-441.

Caldwell, J. (1985). A new look at the old informal reading inventory. The Reading Teacher, 39, 168-173.

Carr, E., Dewitz, P., & Patberg, J. (1989). Using cloze for inference training with expository text. The Reading Teacher, 42, 380-385.

Cromer, W. (1970). The difference model: A new explanation for some reading difficulties. Journal of Educational Psychology, 61, 471-483.

Dale, E., & Chall, J. S. (1948). A formula for predicting readability. Educational Research Bulletin, 27, 11-20, 37-54.

D'Angelo, K., & Mahlios, M. (1983). Insertion and omission miscues of good and poor readers. The Reading Teacher, 36, 778-82.

Davey, B. (1983). Think aloud: Modeling the cognitive processes of reading comprehension. Journal of Reading, 27, 44-47.

Davey, B. (1987). Postpassage questions: Task and reader effects on comprehension and metacomprehension processes. Journal of Reading Behavior, 19, 261-285.

Davey, B. (1988a). Factors affecting the difficulty of reading comprehension items for successful and unsuccessful readers. Journal of Experimental Education, 56, 67-76.

Davey, B. (1988b). The nature of response errors for good and poor readers when permitted to reinspect text during question-answering. American Educational Research Journal, 3, 399-414.

Davey, B. (1989). Assessing comprehension: Selected interactions of task and reader. The Reading Teacher, 42, 694-697.

Davey, B., & LaSasso, C. J. (1984). The interaction of reader and text factors in the assessment of reading comprehension. Journal of Experimental Education, 52, 199-206.

Davey, B., & Macready, G. B. (1985). Prerequisite relations among inference tasks for good and poor readers. Journal of Educational Psychology, 77, 539-552.

Davison, A., & Kantor, R. N. (1982). On the failure of readability formulas to define readable texts: A case study from adaptions. Reading Research Quarterly, 17, 187-209.

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

Duffelmeyer, F. A. (1983). The effect of grade level on cloze test scores. Journal of Reading, 26, 436-441.

Durkin, D. (1989). Teaching them to read. Boston: Allyn & Bacon.

Dyer, P. A., & Lowery, L. F. (1988). Thinking about science: Focus on content reading, (Levels B-F). Cleveland, OH: Modern Curriculum Press.

Englert, C. S., & Hiebert, E. (1984). Children's developing awareness of text structure in expository materials. Journal of Educational Psychology, 76, 65-74.

Erickson, L. G., Stahl, S. A., & Rinehart, S. D. (1985). Metacognitive abilities of above and below average readers: Effects of conceptual tempo, passage level and error type on error detection. Journal of Reading Behavior, 17, 235-252.

Fairbanks, G. (1937). The relation between eye-movements and voice in the oral reading of good and poor silent readers. Psychological Monographs, 48, 78-107.

Farris, P. J., & Triplett, D. (1986). Teacher-student interaction in critical thinking development: Does reading level make a difference? Ohio Reading Teacher, 21, 21-25.

Fitzgerald, J. (1983). Helping readers gain self-control over reading comprehension. The Reading Teacher, 37, 249-253.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. American Psychologist, 32, 906-911.

Flavell, J. H. (1981). Cognitive monitoring. In W. P. Dickson (Ed.), Children's oral communication skills (pp. 35-60). New York: Academic Press.

Flesch, R. F. (1948). A new readability yardstick. Journal of Applied Psychology, 32, 221-233.

Freebody, P., & Anderson, R. C. (1983). Effects on text comprehension of differing proportions and locations of difficult vocabulary. Journal of Reading Behavior, 15, 19-39.

Fry, E. B. (1968). A readability formula that saves time. Journal of Reading, 11, 513-516, 575-578.

Gambrell, L. B., & Bales, R. J. (1986). Mental imagery and the comprehension-monitoring performance of fourth- and fifth-grade poor readers. Reading Research Quarterly, 21, 454-464.

Gambrell, L. B., Wilson, R. M., & Gantt, W. N. (1981). Classroom observations of task-attending behaviors of good and poor readers. Journal of Educational Research, 74, 400-404.

Garner, R. (1980). Monitoring of understanding: An investigation of good and poor readers' awareness of induced miscomprehension of text. Journal of Reading Behavior, 12, 55-63.

Garner, R. (1981). Monitoring of passage inconsistency among poor comprehenders: A preliminary test of the "piecemeal processing" explanation. Journal of Educational Research, 74, 159-162.

Garner, R. (1987). Metacognition and reading comprehension. Norwood, NJ: Ablex.

Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. Review of Educational Research, 60, 517-529.

Garner, R., & Anderson, J. (1982). Monitoring of understanding research: Inquiry, directions, methodological dilemmas. Journal of Experimental Education, 50, 70-76.

Garner, R., & Gillingham, M. G. (1987). Students' knowledge of text structure. Journal of Reading Behavior, 19, 247-259.

Garner, R., & Kraus, C. (1981-1982). Good and poor comprehender differences in knowing and regulating reading behaviors. Educational Research Quarterly, 6, 5-12.

Garner, R., & Reis, R. (1981). Monitoring and resolving comprehension obstacles: An investigation of spontaneous text lookbacks among upper-grade good and poor comprehenders. Reading Research Quarterly, 16, 569-582.

Garner, R., & Taylor, N. (1982). Monitoring of understanding: An investigation of attentional assistance needs at different grade and reading proficiency levels. Reading Psychology, 3, 1-6.

Gentile, L. M., & McMillan, M. M. (1987). Stress and reading difficulties: Teaching students self regulating skills. The Reading Teacher, 41, 170-178.

Gillet, J. W., & Temple, C. (1990) Understanding reading problems: Assessment and instruction, (3rd ed.). New York: Harper Collins.

Goswami, U., & Bryant, P. (1989). The interpretation of studies using the reading level design. Journal of Reading Behavior, 21, 413-424.

Grabe, M., & Mann, S. (1984). A technique for the assessment and training of comprehension monitoring skills. Journal of Reading Behavior, 16, 136-144.

Guthrie, J. T. (1981). Reading interests. The Reading Teacher, 34, 984-986.

Guthrie, J. T. (1983). Children's reasons for success and failure. The Reading Teacher, 36, 478-480.

Hahn, A. L. (1984). Assessing and extending comprehension: Monitoring strategies in the classroom. Reading Horizon, 24, 231-237.

Hansell, T. S. (1976). Readability, syntactic transformations, and generative semantics. Journal of Reading, 19, 557-562.

Hieronymous, A. N., Hoover, H. D., & Lindquist, E. F. (1986). Iowa tests of basic skills. Chicago: Riverside.

Hoffman, J. V., O'Neal, S. F., Kastler, L. A., Clements, R. O., Segel, K. W., & Nash, M. F. (1984). Guided oral reading and miscue focused verbal feedback in second-grade classrooms. Reading Research Quarterly, 19, 367-384.

Holmes, B. C., & Rosser, N. L. (1987). Five ways to assess readers' prior knowledge. The Reading Teacher, 40, 646-649.

Horowitz, R. (1985). Text patterns: Part 1. Journal of Reading, 28, 534-541.

Johnston, P. (1983). Reading comprehension assessment: A cognitive basis. Newark, DE: International Reading Association.

Jones, M. B., & Pikulski, E. C. (1974). Cloze for the classroom. Journal of Reading, 17, 432-438.

Kavale, K., & Schreiner, R. (1979). The reading processes of above average and average readers: A comparison of the use of reasoning strategies in responding to standardized comprehension measures. Reading Research Quarterly, 15, 102-128.

Kibby, M. W. (1979). Passage readability affects the oral reading strategies of disabled readers. The Reading Teacher, 32, 390-396.

Klare, G. R. (1984). Readability. In P. D. Pearson (Ed.), Handbook of reading research (pp. 681-744). New York: Longman.

Kletzien, S. B. (1991). Strategy use by good and poor comprehenders reading expository text of differing levels. Reading Research Quarterly, 26, 67-86.

Lapp, D., & Flood, J. (1984). Promoting reading comprehension: Instruction which insures continuous reader growth. In J. Flood (Ed.), Promoting reading comprehension pp. 273-288). Newark DE: International Reading Association.

Lapp, D., Flood, J., & Farnan, N. (1989). Content area reading and learning: Instructional strategies. Englewood Cliffs, NJ: Prentice-Hall.

Latvala, C. J. (1989). The effects of text difficulty on average fifth-grade readers' comprehension monitoring and reading comprehension. Unpublished doctoral dissertation, University of Maryland, College Park.

Leslie, L. (1980). The use of graphic and contextual information by average and below-average readers. Journal of Reading Behavior, 12, 139-149.

Leslie, L., & Caldwell, J. (1990). Qualitative reading inventory. Glenview, IL: Scott, Foresman.

Leslie, L., & Osol, P. (1978). Changes in oral reading strategies as a function of quantities of miscues. Journal of Reading Behavior, 10, 442-444.

Leu, D. J. (1981-1982). Oral reading analysis: A critical review of research and application. Reading Research Quarterly, 17, 420-437.

Lipson, M. Y., & Wixson, K. K. (1986). Reading disability research: An interactionist perspective. Review of Educational Research, 56, 111-136.

Lively, B. A., & Pressey, S. L. (1923). A method for measuring the "vocabulary burden" of textbooks. Educational Administration and Supervision, 9, 389-398.

Mamen, M., Ferguson, H. B., & Backman, J. E. (1986). No difference represents a significant finding: The logic of the reading level design. Psychological Bulletin, 100, 104-106.

Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. Child Development, 48, 678-683.

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

Marshall, N., & Glock, M. (1978-1979). Comprehension of connected discourse. Reading Research Quarterly, 14, 10-56.

McConaughy, S. H. (1985). Good and poor readers' comprehension of story structure across different input and output modalities. Reading Research Quarterly, 20, n. 2, 219-232.

McCormick, S. (1989). Effects of previews on more skilled and less skilled readers' comprehension of expository text. Journal of Reading Behavior, 21, 219-239.

McGee, L. M., & Richgels, D. J. (1985). Teaching expository text structure to elementary students. The Reading Teacher, 38, 739-748.

McKenna, M. C., & Kear, D. J. (1990). Measuring attitude toward reading: A new tool for teachers. The Reading Teacher, 43, 626-639.

McKeown, J. G., Beck, I. L., Sinatra, G. M., & Loxterman, J. A. (1992). The contribution of prior knowledge and coherent text to comprehension. Reading Research Quarterly, 27, 79-93.

Miles, W. R., & Segel, D. (1929). Clinical observation of eye-movements in the rating of reading ability. Journal of Educational Psychology, 20, 520-529.

Miller, G. E. (1987). The influence of self-instruction on the comprehension monitoring performance of average and above average readers. Journal of Reading Behavior, 19, 303-317.

Mosenthal, P. (1984). The problem of partial specification in translating reading research into practice. Elementary School Journal, 85, 199-227.

Olshavsky, J. E. (1976-1977). Reading as problem solving: An investigation of strategies. Reading Research Quarterly, 12, 654-574.

Olshavsky, J. E. (1978). Comprehension profiles of good and poor readers across material of increasing difficulty. In P. D. Pearson & J. Hansen (Eds.), Reading: Disciplined inquiry in process and practice pp. 73-76). (Twenty-seventh Yearbook). Clemson, SC: The National Reading Conference.

Olson, M. W., & Gee, T. C. (1991). Content reading instruction in the primary grades: Perceptions and strategies. The Reading Teacher, 45, 298-307.

Owings, R. A., Peterson, G. A., Bransford, J. D., Morris, C. D., & Stein, B. S. (1980). Spontaneous monitoring and regulation of learning: A comparison of successful and less successful fith graders. Journal of Educational Psychology, 72, 250-256.

Palinscar, A. S., & Brown, A. L. (1986). Interactive teaching to promote independent learning from text. The Reading Teacher, 39, 771-777.

Paris, S. G., Cross, D. R., & Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. Journal of Educational Psychology, 76, 1239-1252.

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. Contemporary Educational Psychology, 8, 293-316.

Paris, S. G., & Myers, M. II. (1981). Comprehension monitoring, memory and study strategies of good and poor readers. Journal of Reading Behavior, 13, 5-22.

Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. Contemporary Education Psychology, 8, 317-344.

Pearson, P. D., & Johnson, D. D. (1978). Teaching reading comprehension. New York Holt, Rinehart & Winston.

Piccolo, J. (1987). Expository text structure: Teaching and learning strategies. The Reading Teacher, 40, 838-847.

Pikulski, J. J., & Shananhan, T. (1982). Informal reading inventions: A critical analysis. In J. J. Pikulski & T. Shanahan (Eds.), Approaches to the informal evaluation of reading (pp.94-116). Newark, DE: International Reading Association.

Pikulski, J. J., & Tobin, A. W. (1982). The cloze procedure as an informal assessment technique. In J. J. Pikulski & T. Shanahan (Eds.), Approaches to the informal evaluation of reading. Newark, DE: International Reading Association.

Pitts, M. M. (1983). Comprehension monitoring: Definition and practice. Journal of Reading, 26, 516-523.

Rahman, T., & Bisanz, G. L. (1986). Reading ability and use of a story schema in recalling and reconstructing information. Journal of Educational Psychology, 78, n. 5, 323-333.

Random House Achievement Program in Comprehension. (1988). New York: Random House.

Rankin, E., & Culhane, J. (1969). Comparable cloze and multiple-choice test scores. Journal of Reading, 13, 193-198.

Raphael, T. E., Myers, A. D., Tirre, W. C., Freebody, P., & Fritz, M. (1981). The effects of some known sources of reading difficulty on metacomprehension and comprehension. Journal of Reading Behavior, 13, 325-334.

Reader's Digest (1960). Reader's Digest skill builders III. Pleasantville, NY: Reader's Digest.

Recht, D. R., & Leslie, L. (1988). The effects of prior knowledge on good and poor readers' memory for text. Journal of Educational Psychology, 80, 16-20.

Reis, R., & Spekman, N. J. (1983). The detection of readers-based versus text-based inconsistencies and the effects of direct training of comprehension monitoring among upper-grade poor comprehenders. Journal of Reading Behavior, 15, 49-60.

Richards, J. (1978). Classroom language: What sorts? London: Allen & Unwin.

Richgels, D. J. (1982). Schema theory, linguistic theory, and representations of reading comprehension. Journal of Educational Research, 76, 54-62.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension (pp. 33-58). Hillsdale, NJ: Lawrence Erlbaum.

Ryan, E. B. (1981). Identifying and remediating failures in reading comprehension: Towards an instrumental approach for poor comprehenders. In G. E. MacKinnon & T. Q. Waller (Eds.), Reading research advances in theory and practice (pp. 223-261). New York: Academic Press.

Sawyer, M. H. (1991). A review of research in revising instructional text. Journal of Reading Behavior, 23, 307-333.

Short, E. J., & Ryan, E. B. (1984). Metacognitive differences between skilled and less skilled readers: Remediating deficits through story grammar and attribution training. Journal of Educational Psychology, 76, 225-235.

Siedow, M. D., & Fox, B. J. (1984). Effects of training on good and poor readers' use of top level structure. Reading World, 23, 340-346.

Singer, H., & Donlan, D. (1989). Reading and learning from text, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum.

Singer, M. H. (1982). Reading disability research: A misguided search for differences. In M. H. Singer (Ed.), Competent reader, disabled reader (pp. 39-53). Hillsdale, NJ: Lawrence Erlbaum.

Smith, F. (1988). Joining the literacy club. Portsmouth, NH: Heinemann.

Smith, H. K. (1967). The responses of good and poor readers when asked to read for different purposes. Reading Research Quarterly, 3, 53-84.

Spache, G. (1953). A new readability formula for primary-grade reading materials. Elementary School Journal, 53, 410-413.

Stahl, S. A., Rinehart, S. D., & Erickson, L. G. (1985-1986). Detection of inconsistencies by above and below average reflective and impulsive sixth graders. Journal of Educational Research, 79, 185-189.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. Reading Research Quarterly, 21, 360-407.

Stanovich, K. E., Nathan, R. G., & Vala-Rossi, M. (1986). Developmental changes in the cognitive correlates of reading ability and the developmental lag hypothesis. Reading Research Quarterly, 21, 267-283.

Steele, J. L. (1985). Recall and comprehension: The interactive relationship of text and reader. Unpublished doctoral dissertation, University of Virginia, Charlottesville.

Swaby, B. E. R. (1989). Diagnosis and correction of reading difficulties. Boston: Allyn & Bacon.

Taylor, B., Harris, L. A., & Pearson, P. D. (1988). Reading difficulties: Instruction and assessment. New York: Random House.

Taylor, W. (1953). Cloze procedure: A new tool for measuring readability. Journalism Quarterly, 30, 415-433.

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

Tregaskes, M. R., & Daines, D. (1989). Effects of metacognitive strategies on reading comprehension. Reading Research and Instruction, 29, 52-60.

Treiman, R., & Hirsch-Pasek, K. (1985). Are there qualitative differences in reading behavior between dyslexics and normal readers. Memory and Cognition, 13, 357-364.

Wagoner, S. A. (1983). Comprehension monitoring: What it is and what we know about it. Reading Research Quarterly, 18, 328-346.

Weiner, B. (1974). Achievement motivation as conceptualized by an attribution theory. Morristown, NJ: General Learning Press.

Williamson, L. E., & Young, F. (1974). The IRI and RMI diagnostic concepts should be synthesized. Journal of Reading Behavior, 6, 183-194.

Wilson, P., & Anderson, R. (1986). What they don't know will hurt them: The role of prior knowledge in comprehension. In J. Orasanu (Ed.), Reading comprehension: From research to practice (pp. 31-48). Hillsdale, NJ: Lawrence Erlbaum.

Wilson, R. M., & Gambrell, L. B. (1988). Reading comprehension in the elementary school. Newton, MA: Allyn & Bacon.

Winograd, P., & Johnston, P. (1982). Comprehension monitoring and the error detection paradigm. Journal of Reading Behavior, 24, 61-76.

Winograd, P., & Paris, S. G. (1989). A cognitive and motivational agenda for reading instruction. Educational Leadership, 46, 30-36.

Wixson, K. K., & Lipson, M. Y. (in press). Reading diagnosis and remediation. Glenview, IL: Scott, Foresman.

Wong, B. Y. L. (1986). Metacognition and special education: A review of a view. The Journal of Special Education, 20, 9-29.

Yochum, N. (1991). Children's learning from informational text: The relationship between prior knowledge and text structure. Journal of Reading Behavior, 23, 87-103.

Zabrucky, K., & Moore, D. (1989). Children's ability to use three standards to evaluate their comprehension of text. Reading Research Quarterly, 24, 336-352.

Zabrucky, K., Moore, D., & Ratner, H. H. (1985). Comprehension monitoring in good and poor readers: Evidence from verbal and performance measures. Reading Psychology, 6, 241-250.

Zabrucky, K., & Ratner, H. H. (1986). Children's comprenension monitoring and recall of inconsistent stories. Child Development, 57, 1401-1418.

APPENDIX A

Sample Cloze Passage

Sample Cloze Passage

The Flying Mammal

When you think of ghosts and witches, do you also think of bats? Many people do! Maybe _____ is because bats live _____ dark places and are _____ at night.

Most bats _____ harmless, but they're feared _____ few people know much _____ them.

How many bats _____ on earth? Billions and _____! There are more bats _____ any other mammal except _____ rodents, and bats live _____ everywhere! Scientists have found _____ 850 species, or kinds, _____ bats in almost all _____ of the world. But _____ bats live in the _____.

What do bats eat? _____ North American bats _____ on insects. But there _____ fruit bats that eat _____, fruits, nectar, and pollen. _____ bats feast on other _____ blood. Others eat fish, _____ mammals, and birds.

Bats' _____ habits may be harmful _____ people. For example, in _____ tropics, certain fruit bats _____ large losses to farmers. _____ called vampire bats may

_____ the disease of rabies _____ they bite livestock. But _____ often than not, bats _____ helpful. Some of them _____ flowers. Others eat large _____ of insect pests.

Can_____ really fly? No matter _____ they live, all bats _____ one thing in common. _____ are the only mammals _____ fly.

A bat's wings _____ thin, leathery, elastic pieces _____ skin that stretch between _____ body and its arms _____ fingers. You can get _____ idea of what this _____ if like. Spread the _____ and first finger of _____ hand. Use the fingers _____ your other hand to _____ the double layer of _____ that is stretched thin. That skin is a little like a bat's wing, only the bat's wing is much bigger and stretches from all its fingers to its body, back and tail.

APPENDIX B

Parent Permission Letter--Return Required

Parent Permission Letter--Return Required

February 15, 1991

To the Parents or Guardians of (Student Name),

I am conducting a research study on reading in connection with my Doctorate in Education at the University of Northern Iowa. I have received permission from Waterloo Community Schools and from (Principal's Name) to conduct part of my research at (School).

For this research, a small group of children will participate in a reading session of about one hour's duration. The time will be arranged with the teacher so that your child's schedule is disrupted as little as possible. It will take place during the regular school day. With your permission, I would like for your child to participate. All participants will remain anonymous.

What I expect to learn from this study will help improve our knowledge about reading instruction. The results will be shared with the Waterloo Schools so that your child and others can benefit from my findings. I hope you will grant permission for your child to participate in this reading study. Will you please sign below indicating whether or not you grant permission for your child to participate and return this letter in the enclosed envelope as soon as possible.

Thank you very much,

Maelou Baxter
Doctoral Student
University of Northern Iowa

_____ Yes, my child may participate

_____ No, my child may not participate

Child's Name: _____

Parent's or Guardian's
Signature:_____

APPENDIX C

Error Detection Passages

APPENDIX D

Instructions to Students--Error Detection

Instructions to Students--Error Detection

Hi, I'm Mrs. Baxter. I'm a student at UNI, and I'm doing research on reading. Who knows what research is? (Discuss: let them know they're part of a research project).

Today, I have some reading for you to do. In the six passages I'm going to give you today, there are some mistakes. Each mistake is just one word, and that word is a real word, but it's a word that doesn't make sense in the paragraph.

I want you to first read the title of the story, then read each paragraph and decide whether there is a word that doesn't make sense in the paragraph. If there is, mark yes , and go back and underline the word. Then go on to the next paragraph. If your answer is no, mark no, and go on to the next paragraph. Don't try to decide whether the information is correct, just look for single words that don't seem to go with what the passage is about.

Let's do one together before you begin.
(TRANSPARENCY OF PRACTICE ERROR DETECTION PASSAGE - Demonstrate)

Do all the pages and do the passages in order by the numbers at the top. You may begin when you get your papers. You are not going to be timed. If you finish before everyone else is finished, please read quietly until everyone is through.

As soon as you get your papers, please write your name at the top of the first page of each passage. Remember, you are part of a research project that will help a lot of teachers learn more about teaching reading, so do your best on each passage.

APPENDIX E

Sample Passage--Error Detection

Sample Passage--Error Detection

The Story of a Treehouse

No matter how lonely a tree looks, it is never alone. It always has company. Plants and animals live in it, on it, under it, and around it.

Is there a problem with this paragraph?

___yes___no.

If there is a problem, go back and underline the word that seems to cause the problem.

At the top of a tree a big ball of leaves signals a grey squirrel's summer home. Among the lower branches sits the nest of a wood thrush. This bird, like many others, feeds on insects that harm trees by eating their telephones. One such insect is the tent caterpillar. Its filmy tent covers bushes and trees in the springtime.

Is there a problem with this paragraph?

___yes___no.

If there is a problem, go back and underline the word that seems to cause the problem.

Some tree creatures, like the wood thrush, help the tree. But others harm it. Beetles make small holes in the tree. They dig tunnels under the bark as they search for spoons and a place to lay their eggs. So that it can live and grow, a tree must get food to all its parts. Too many holes and tunnels cut off the tree's food supply. Then the tree starves.

Is there a problem with this paragraph?

___yes___no.

If there is a problem, go back and underline the word that seems to cause the problem.

Like other green plants, the tree uses its leaves to make food. If too many leaves are eaten by caterpillars, the tree cannot make food and again, it starves.

Is there a problem with this paragraph?

___yes___no.

If there is a problem, go back and underline the word that seems to cause the problem.

Some creatures make a dead tree their home. A screech owl might move into the hole left by a fallen pencil. Sometimes a hole is large enough for a family of raccoons, or a bear, looking for a winter home.

Is there a problem with this paragraph?

___yes___no.

If there is a problem, go back and underline the word that seems to cause the problem.

The holes increase in size and number. Soon the dead tree is too weak to stand. In a strong wind, the tree tumbles down.

Is there a problem with this paragraph?

___yes___no.

If there is a problem, go back and underline the word that seems to cause the problem.

The fallen tree is a log. Many creatures are drawn to a log, too. Where the tree's roots were in the ground, there is now a hole. A fox may choose to make this its home. The hollow trunk might become a skunk's home. Snakes, frogs, and toads move in beneath the kitchen, where it is cool and moist.

Is there a problem with this paragraph?

___yes___no.

If there is a problem, go back and underline the word that seems to cause the problem.

APPENDIX F

Parent Permission Letter--No Return Required

Parent Permission Letter--No Return Required

Dear Parent or Guardian,

I am conducting a reading research study in connection with my doctorate in education at the University of Northern Iowa. I have received permission from (School District) and from (Principal's Name) to try out my materials with Longfellow fourth-graders during their regular class time. Your child's name will not be used in connection with this study.

If you do not wish your child to participate, please mark below, sign, and return to school. You do not need to return the form if you have no objection to your child's participation.

I hope you will allow your child to participate so that we will learn more about how children read.

Thank you,

Maelou Baxter
Doctoral Student
University of Northern Iowa

_____No, I do not wish my child to participate in this study

Child's Name: _____

Parent's Signature: _____

APPENDIX G

Individual Students' Scores on the Individual Passages

Individual Students' Scores on the Individual Passages

Stud	*Abil	Stand 1	Stand 2	App 1	App 2	Diff I	Diff 2
1	1	1.00	1.00	.88	1.00	1.00	.75
2	1	1.00	1.00	1.00	.75	1.00	.88
3	1	.00	.50	.75	.50	.43	.63
4	1	1.00	1.00	1.00	1.00	1.00	1.00
5	1	.88	1.00	.88	1.00	.57	.88
6	1	.75	.83	.88	1.00	.86	.88
7	1	1.00	1.00	.88	.88	.71	1.00
8	1	1.00	1.00	1.00	1.00	1.00	1.00
9	1	1.00	1.00	.88	1.00	.86	1.00
10	1	1.00	1.00	1.00	1.00	.71	.75
11	1	1.00	1.00	.75	1.00	1.00	1.00
12	1	.75	1.00	.50	.88	.71	1.00
13	1	.88	1.00	1.00	.88	.43	1.00
14	1	1.00	.67	.88	.88	.86	.88
15	1	.88	1.00	1.00	1.00	.57	.88
16	1	.88	1.00	.63	.88	.86	.88
17	2	1.00	1.00	1.00	1.00	.86	1.00
18	2	.67	1.00	.75	1.00	.57	.88
19	2	.63	.83	.86	.88	.63	1.00
20	2	.83	.17	.57	.50	.50	.43
21	2	.75	1.00	.86	.88	.63	.86
22	2	.88	.83	.71	.88	.88	.86
23	2	1.00	.83	.71	1.00	1.00	.86
24	2	1.00	1.00	1.00	.75	1.00	1.00
25	2	.75	.67	.71	.25	.88	.57
26	2	.50	.50	.43	.50	.63	.57
27	2	.75	.83	.57	.88	.75	.71
28	2	.63	1.00	.71	.88	.75	.86
29	2	1.00	.83	.86	.63	.88	1.00
30	2	.75	.83	.86	.75	.88	.57
31	2	.88	.83	1.00	.88	1.00	1.00
32	2	.88	.50	.86	.75	1.00	.57
33	2	.75	.50	.86	.88	.88	.57
34	2	.75	.83	.86	.88	.75	.86
35	2	.63	.83	.57	1.00	.38	.71
36	2	1.00	.83	.71	.88	1.00	.86

*Ability. Above average readers (1); Below-average readers (2)