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.

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.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Fourth Graders' Literal and Inferential Reading Comprehension: Effects of Readability and Answer Format

Laura Brueggeman Green

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

2000

Program Authorized to Offer Degree: Speech and Hearing Sciences

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to University Microfilms, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, to which the author has granted "the right to reproduce and sell (a) copies of the manuscript in microfilm and/or (b) printed copies of the manuscript made from microform."

Signature: Laura B Green

Date: 4/26/00

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

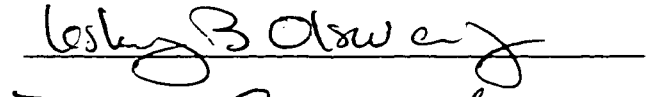
Laura Brueggeman Green

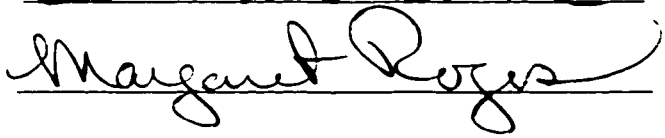
and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of Supervisory Committee

Robert L. Carpenter

Reading Committee:





Date: 4/26/00

University of Washington

Abstract

Fourth Graders' Literal and Inferential Reading Comprehension: Effects of Readability and Answer Format

Laura Brueggeman Green

Chairperson of the Supervisory Committee:
Professor Robert L. Carpenter
Speech and Hearing Sciences

The goal of reading is to gain meaning from text. Subsequently, reading comprehension skills become a much greater focus in third or fourth grade as children begin reading text to learn. A text's meaning is a combination of the explicit, literal meanings of the sentences themselves, as well as the inferential meanings that can be uniquely generated by the reader. Inferential comprehension is critical to reading success. Assessment of inferential reading comprehension in the classroom is often completed in the context of passage-reading tasks (i.e. read a passage and answer questions that follow). Many variables can influence and potentially confound this type of assessment. The purpose of this study was to examine the effects of two such variables, text readability level and the use of different answer formats, on fourth-grade readers' ability to gain literal and inferential meaning during reading.

Sixty normally-achieving fourth graders participated in the study. Tasks were comprised of passages and items that varied as to *readability level* (below-grade-level vs. grade-appropriate), and items that varied as to *answer format* (multiple-choice vs. short-answer), and *question type* (literal vs. inferential). Once the participants completed all three tasks, the items were scored and reorganized into eight separate subscales used for comparison in this study.

A 3-factor repeated-measures analysis of variance was used to examine effects of readability and answer format on literal and inferential comprehension performance. Participants performed consistently *better* at *below-grade-level readability levels* on both literal and inferential comprehension and in both answer formats. Moreover, there were differential effects of answer format on literal comprehension performance at different readability levels. On *below-grade-level literal subscales*, participant performance on literal questions was significantly better in the *multiple-choice* format; however, at a *grade-appropriate readability level*, participant performance on *literal* questions was significantly better in the *short-answer* format. Answer format did not significantly affect performance on inferential questions.

These results indicate that both readability and answer format influence assessment of literal and inferential comprehension performance, and should be taken into account if accurate evaluation is to take place.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to acknowledge Project II of the University of Washington Multidisciplinary Learning Disabilities Center for allowing the inclusion of my study in the research conducted during the 1998-1999 school year.

I would like to express my appreciation to the members of my committee, Lesley Olswang, Deborah McCutchen, and Margaret Rogers for their valuable feedback in the development of my dissertation, and to my advisor, Bob Carpenter, for his editorial expertise and support throughout this project.

I'd also like to thank Tasha Beretvas for being a great friend and statistical consultant, and all of my friends, especially Ellen Hickey, Geralyn Timler, Brenda Long, and Megan Lacey for their friendship, encouragement, and willingness to listen.

And most importantly, I am grateful to my husband Tim, without whose love, support, and understanding, I would not be where I am today. Thank you for being who you are.

Chapter 1

Introduction

When children move from second into third and fourth grade, they must make the shift from learning to read (decoding) to reading to learn (comprehension). Subsequently, the focus of reading instruction changes in the upper elementary grades. Reading is a meaning-acquisition process (Thorndike, 1917). Given that the ultimate goal of reading is to gain meaning from print, reading comprehension skills become a much greater focus in third and fourth grade classrooms and beyond. Reading comprehension is a process with many components that work together in synergy to generate meaning in the mind of the reader. While a great deal is known about the requisite skills needed for decoding (i.e. phonological awareness, orthographic awareness) and variables that might influence decoding success (i.e. word length and complexity), there is less known about reading comprehension. Better understanding of reading comprehension begins by considering different types of comprehension (i.e. literal and inferential) and factors that affect children's ability to read text for meaning.

The meaning of a passage of text is not straightforward. It is a combination of the explicit, literal meanings of the words and sentences, as well as the inferred (implicit) meanings that can be uniquely generated by the reader. Given that meaning is not considered to be given solely in the text, but is mentally constructed by readers during the reading process (Maria, 1990), the total message within the written discourse is dependent upon the reader applying additional knowledge and "reading between the lines." Successful reading comprehension involves gaining both literal meaning as well as deeper, implicit meaning.

The key to the generation of implicit, higher level meaning involves the reader's ability to make inferences. The ability to generate inferences is an essential skill that determines, to a large extent, the degree to which a passage will be understood (Casteel, 1993; Omanson, Warren, & Trabasso, 1978; Zabrusky, 1986). Research in education today still strives for a better understanding of this higher level of reading comprehension that is so critical for academic success. The purpose of this study is to examine influential factors in fourth graders' generation of literal and inferential meaning during reading tasks.

Reading Comprehension

Faced with written discourse, which is comprised of words and sentences, the reader must formulate a coherent interpretation of the information as a whole. This requires application of knowledge to generate both explicit and implicit meaning. This total meaning, or "gestalt," can be characterized partially by the inferences the skilled reader must make to connect the meanings of the various sentences in a sensible way (Sanford and Garrod, 1981). An inference is created when the reader activates information that is evoked by, yet goes beyond, the information provided specifically in the text. From this theoretical

perspective, overall reading comprehension is dependent on the ability to formulate appropriate inferences as well as appreciate explicit, literal textual meanings. For example, in order to comprehend the following short passage, both literal and inferential comprehension must occur.

> *Carol was fed up with her job. Customers were rude, the chef was impossibly demanding, and the manager had made a pass at her just that day. The last straw came when a rude man at one of her tables complained that the spaghetti she had just served was cold. He becomes louder and nastier and, without thinking of the consequences, she picked up the plate of spaghetti, and raised it above the rude man's head.*
> *(Murray, Klin & Myers, 1993)*

Literal meaning must be generated such as appreciating that Carol didn't like her job or the chef was overly demanding. Beyond this, inferences must be made for complete understanding. The passage never states that Carol is a waitress, that Carol became very angry, or that she dumped the spaghetti on the man's head. All of these conclusions must be inferred for the total meaning of the text to be appreciated. A number of processes must come together that allow literal and inferential comprehension of written discourse.

Figure 1 (based on Mitchell, 1982) illustrates this "process" of reading comprehension as a function of working memory (represented by rectangles), different cognitive operations (represented by circles), influential variables (represented by triangles) and long-term knowledge bases (represented by octagons). Working memory is the set of resources that allows information to be stored and analyzed simultaneously, cognitive operations involve active, conscious processing, influential variables incorporate factors intrinsic (e.g. reader's objectives, intelligence) and extrinsic (e.g. readability) to the reader, and long-term knowledge bases are storehouses of information that are tapped to help make sense of what is read.

While the above components of reading comprehension work in an ongoing and simultaneous manner throughout the process, the model necessitates their discussion in a somewhat linear progression. First, words must be decoded through the analysis of the visual features of the letters, and then recognized. Utilizing short-term, and then working memory, the words are stored, matched to the semantic knowledge base in long-term memory, and identified. Then, propositions (idea units) are constructed. Utilizing the storage and processing of working memory, the reader continues to decode, syntactic parsing of word sequences occurs, and propositions continue to be formed. The reader has at this stage constructed the "surface code," which is the first level of attaching meaning to text. Then, the reader utilizes world knowledge stored in long-term memory to link these ideas, help maintain coherence, and establish the level of the "textbase." Lastly, the reader makes necessary inferences and builds a mental model or "situation model" of the text for ultimate comprehension (Kintsch, 1978). This is the "total" meaning, or mental representation, of the people, setting, actions, and events that are mentioned "explicitly" in the

clauses and sentences, and that are augmented "inferentially" by application of world knowledge (Graesser, Singer, & Trabasso, 1994).

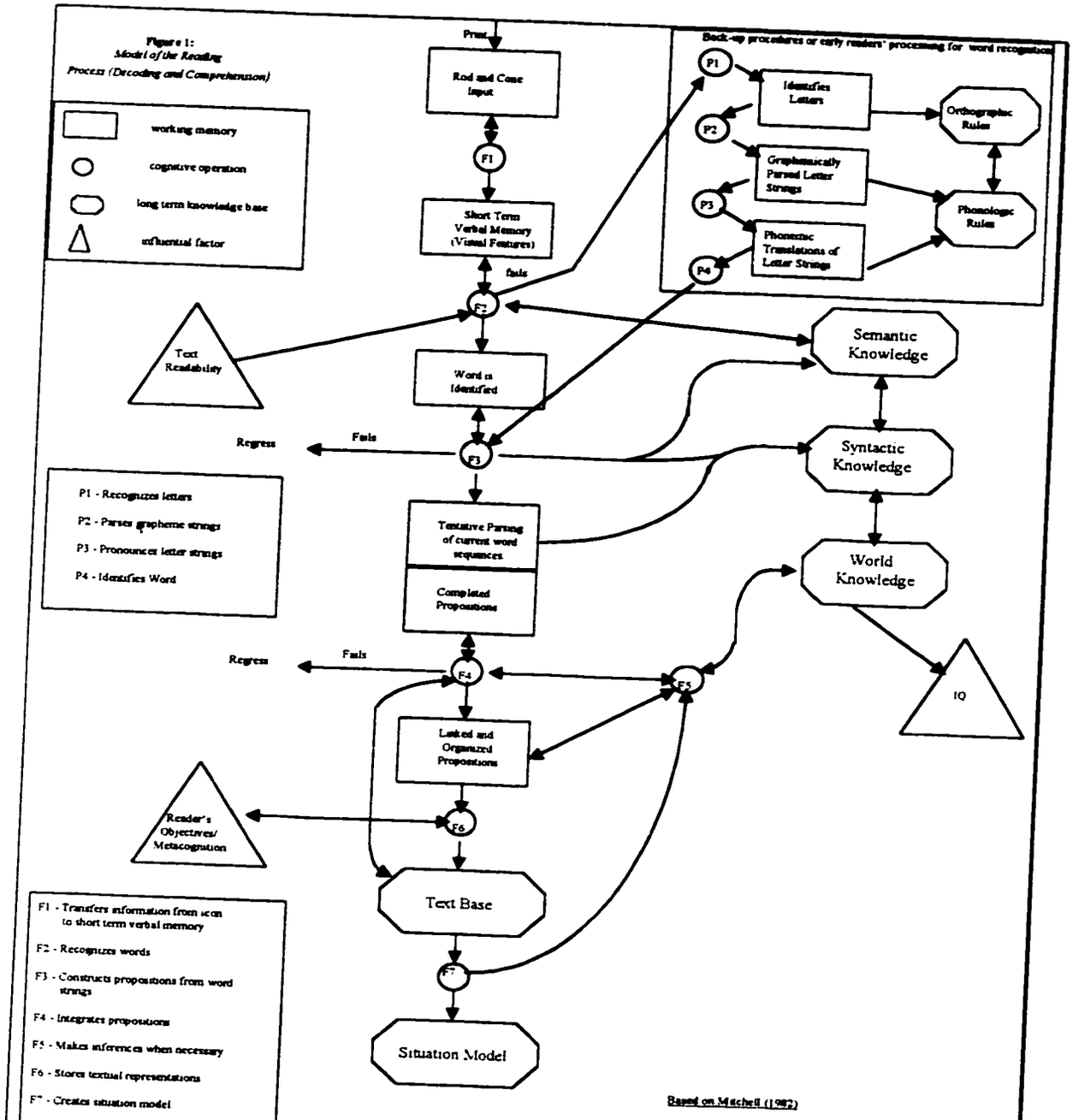


Figure 1: Model of the Reading Comprehension Process

Inferential Reading Comprehension: Children in the Classroom

Reading for meaning in the classroom is highly dependent upon students' ability to successfully perceive both literal and inferential meanings. In almost every classroom, students are asked to read various types of written discourse in the form of basal readers, childrens' literature, subject area textbooks, reference materials and weekly readers (Lipson & Wixon, 1997). By third or fourth grade, reading comprehension is a skill students rely on daily, and inferential comprehension provides the complete understanding of material that cannot be gleaned from literal meaning alone.

Given its contribution to successful reading comprehension, the ability to make inferences during reading has been studied across all ages during the last several decades. Developmental studies have not been as common as studies with adults, and reading comprehension contexts have not been utilized as often as listening comprehension contexts. However, the reading comprehension studies undertaken have demonstrated that children make inferences relatively early (by age six), but become much more comfortable with inferential processing during their school years, after they are seven or eight years old (Blanchowitz, 1986; Westby, 1984). Most school-age readers have more difficulty with inferential comprehension as opposed to factual recall (Hansen, 1981; Holmes, 1984); however, skilled or average readers have performed more successfully than less skilled readers on inferential questions in reading tasks (Holmes, 1984; Schacter, 1980; Schreiner & Shannon, 1980; Wilson, 1979). The two most important findings with respect to reading assessment for school-age children are that inferential comprehension is consistently a more difficult element than literal comprehension in the reading process, and that there are students who have specific difficulty with this skill.

Beyond these general conclusions, little is known about inferential comprehension in functional classroom contexts with typical classroom reading material, and less is known about factors that can influence successful inferential comprehension or how such factors should be considered when reading comprehension is evaluated in school age children.

Assessment of Inferential Reading Comprehension: Research and Practice

Given that inferential comprehension is critical to reading success and is apparently more difficult than literal comprehension, teachers and researchers must be able to accurately and differentially assess inferential reading skill in functional contexts. Students may need to be taught specifically how to think inferentially during reading, and some form of assessment is necessary to determine how this instruction should proceed. While the research literature has provided many interesting insights into inferential abilities in school age children, very few studies have actually examined these skills in functional reading contexts. The studies that did examine reading included sentence reading tasks and, occasionally, narrative reading tasks. The research was conducted with subjects in one-on-one situations and with the examiners asking questions and the children answering orally. No studies have been conducted in the classroom setting with materials that students are typically exposed to in their daily language arts and

reading activities. In order to better understand inferential comprehension in the context of the classroom, research should be conducted in this setting.

In addition to this issue, there are several problems with the assessment of reading comprehension in the classroom. First, inferential comprehension skills are typically not examined separately from or relative to literal comprehension skills. In the context of standardized reading achievement tests such as the Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1989), students earn an overall "comprehension" score only. If students perform poorly on the test and teachers don't have information as to the types of questions students are answering (literal vs. inferential), the nature of reading comprehension difficulties remains enigmatic (i.e. Is literal comprehension impaired? Is inferential comprehension impaired? Are both types impaired?). Only when both literal and inferential comprehension skills are evaluated and examined relative to one another, can a true sense of inferential reading comprehension ability be attained.

A second problem with classroom reading comprehension assessments is that, given the nature of the tasks often utilized, inferential comprehension skills may not be assessed validly. Accurate evaluation is dependent upon the knowledge that inferential comprehension is indeed the construct being measured. Yet there are many variables that can potentially influence comprehension during reading tasks and confound accurate assessment. These variables may be categorized as those within the reader (i.e. world knowledge, syntactic knowledge, motivation) and as those that are extrinsic to the reader (i.e. text readability, task demands). With regard to extrinsic variables, assessment must take into account text and task variables before comprehension performance can be accurately evaluated. Students are presented with an enormous array of materials and tasks during their development as readers, and the context of these printed materials will have a significant impact on both how much and what students understand and learn from a text (Lipson & Wixon, 1997). Subsequently, research in inferential comprehension assessment must isolate relevant contextual variables and sort out their exact contribution to comprehension performance. Two such influential contextual factors are text readability and the answer format (i.e. multiple-choice, short-answer) utilized in reading comprehension tasks. Both will be examined in this study as to the role they may play in the assessment of literal and inferential reading comprehension.

Readability

Readability may be thought of as the ease with which a text may be comprehended, and it may be influenced by factors such as word length, word frequency, number of syllables per word, and sentence length (Samuels, 1982). Many formulas have been created to determine text readability and this concept has been used to estimate the amount of difficulty students might encounter when reading a text.

The instructional implications of text readability are great. The ease with which a student can decode a passage greatly determines comprehension ability. There are several assessment inventories that manipulate readability in order to evaluate a student's "reading level." A commonly used taxonomy to scale readability incorporates three readability levels. The independent level is the level at which a student can read successfully *without assistance*. This reading level should be utilized for students' free reading and for tasks upon which the reader is expected to perform independently. Materials at an independent level are also suggested for reading strategy instruction or fluency practice. This allows the reader to learn and practice a strategy on relatively easy text before transferring to more challenging material (Leslie & Caldwell, 1995). The next level, the instructional level, is that at which a student can read *with assistance* from a teacher. Materials written at this level should be chosen for reading and content area instruction. This level assumes that the teacher will introduce words and concepts, which are likely to be unfamiliar to the readers, and will provide appropriate background knowledge necessary for understanding the material. The final level, the frustration level, is that at which the student is completely *unable to read the material* with adequate word identification or comprehension. Teachers should avoid materials at this level (Leslie & Caldwell, 1995).

The implications for uncontrolled readability levels when assessing inferential comprehension are clear. If students have difficulty with decoding or word recognition and are unable to accurately read the material, comprehension errors may seem like an inferential thinking problem, but may instead be a more basic decoding problem. If readability is too difficult, they will not be able glean the necessary information to form inferences and successfully comprehend the text. Even good readers may have to devote additional cognitive resources to decoding text at grade-appropriate or higher readability levels. The descriptions of the reading levels above, and the fact that inferential comprehension is more difficult than literal comprehension, suggest that students may need to learn inferential reading strategies with text at an independent level. Yet, given the variance in student abilities and the instructional nature of regular classroom language arts activities, this may not be the case for many students. If decoding utilizes cognitive resources that may affect students' ability to make inferences, the influence of readability level on inferential performance must be investigated and better understood.

The influence of readability level on a higher-level process such as inference generation is further supported by reading models such as the LaBerge-Samuels model of reading (LaBerge & Samuels, 1974). They suggested that the development of rapid, automatic processing of words may be required for attention to be focused on comprehension processes when reading. Subsequently, a lack of automaticity in decoding can have a negative effect on comprehension. The model assumes that attention (e.g. cognitive processing resources) is used in decoding and comprehension and, since there is a limit to the amount of attention available for any time period, it is important to have enough attention available to accomplish both of these reading tasks (Samuels, 1987). When a task requires greater information processing

demands, cognitive resources will be tapped at more basic, lower levels and may not be available for higher level processing. As a result the human brain is often referred to as a "limited capacity processor," and readability may affect its processing performance.

Another similar theoretical issue related to breakdowns in reading processing involves the limited resource known as working memory. Working memory is conceptualized as an arena for mental computation as well as storage. It consists of a pool of operational resources needed to perform decoding and comprehension computations, and storage functions to maintain the intermediate or partial products of comprehension generated on the way to a final interpretation. The capacity of working memory is defined as the total amount of activation, or "fuel," that can be allocated to support information processing and storage concurrently. Working memory has a limited capacity and therefore, when simultaneous processing is being carried out, there is competition for access to resources (Just & Carpenter, 1992). It is presumed, that when processing demands are high and exceed working memory capacity, comprehension will break down. Both the LaBerge-Samuels model of reading and the theory behind working memory suggest that, if readability of a text is too difficult, students may not have resources left for higher-level comprehension processes. Text readability, then, can tax students' "resource capacity," and may be a very influential factor in successful inferential comprehension during reading.

Answer Format

A second critical issue that can influence accurate assessment of inference generation during reading comprehension involves the *manner* in which inferential comprehension is examined. When reading comprehension is assessed in the classroom, students are typically asked to answer questions after reading a passage. The format of these questions can have significant impact on the determination of a student's inferential reading comprehension. Reading comprehension questions usually utilize either a multiple-choice or a short-answer (constructed) response. Each of these formats requires different skill application on the part of the student. For example, in answering multiple-choice questions, students may capitalize on abilities that have little to do with the construct being measured. A response-elimination strategy has been identified, whereby readers eliminate implausible distracters and then guess from the remaining options (Snow, 1980). Also, even without a specific strategy, it is often possible to guess the correct answer. Multiple-choice questions may also often give clues to the meaning of the passage, leading a weak student to a level of "comprehension" which he may not have achieved alone (Stathmann, 1977).

These problems do not arise with constructed-response items, which require an independently generated answer. Contrasts between constructed-response and multiple-choice questions have been examined in various contexts using parallel or "stem-equivalent" questions (e.g. questions with equivalent forms), but requiring different answer types. In general, constructed-response questions have been found to be more "difficult" and "reliable" than stem-equivalent multiple-choice counterparts (Traub & MacRury, 1986). For example, in a study by Martinez (1991), item and test statistics from parallel sets of constructed-

response and multiple-choice questions were compared. Constructed response items were generally more difficult, especially for questions that were harder to answer from a conceptual standpoint. As a result, different answer formats may affect students' abilities to answer inferential questions. For example, if multiple-choice questions are utilized, students' inferential comprehension abilities may be overestimated because the answers suggested in the multiple-choice framework may cause an individual to make an inference that was not originally, independently generated. Until more is known about the direct contribution of answer format to inference generation, there can be no certainty as to what type of evaluative information a question-answering format provides.

Chapter II
Research Questions

Assessment of reading comprehension in the classroom often utilizes question-answering tasks (i.e. reading a passage and answering questions that follow) to make decisions about student reading ability. Readability level and answer format have the potential to influence performance on these tasks. Therefore, the effects of readability and answer format on inferential comprehension performance should be examined in the context of a question-answering paradigm. To better understand the effects of these two variables on inferential comprehension performance, their influence on literal reading comprehension tasks will also be examined. The inclusion of both literal and inferential contexts will allow for comprehensive analysis of the influences of readability and answer format, for examination of differences in the two types of comprehension (i.e. Is inferential comprehension more difficult?), and for isolation of differential effects on inferential comprehension performance (i.e. Does answer format affect inferential comprehension performance differently than literal comprehension performance?).

The specific questions are:

I. How does *readability level* influence literal and inferential comprehension performance on question-answering tasks for fourth grade readers?

It is hypothesized that:

A. If readability *does not* affect *literal* comprehension performance, but *does* affect *inferential* performance, then there is a difference between the two types of comprehension (i.e. inferential comprehension is more resource-demanding) and both types should be included separately in assessment.

B. If inferential performance is better at an *easier* readability level then, once the reading difficulty has been controlled, more accurate inferential processing can take place.

C. If the easier readability level does *not* facilitate the ability to answer inferential questions, then perhaps there are other issues that make inferential comprehension more difficult and could influence accurate assessment.

II. How does *answer format* influence literal and inferential comprehension performance on question-answering tasks for fourth grade readers?

It is hypothesized that:

A. If answer format *does not* affect performance on *literal* tasks, but *does* affect performance on *inferential* tasks, then inferential comprehension may be more difficult and more susceptible to alterations in task demands.

B. If a multiple-choice format (as opposed to a short-answer format) makes it easier to answer inferential questions, then performance may be facilitated by a closed set of responses. An inference may not be made until the multiple choices trigger the appropriate connection of ideas. If children have more difficulty in the short-answer format (where they must independently formulate an inferential answer), this might indicate that inferences are not necessarily being made readily during reading. If so, this latter format may provide a more appropriate representation of actual inferential comprehension skill.

Chapter III

Method

Participants

Sixty fourth-graders participated in the study during September and October, 1998. Students at this grade level were selected on the premise that inferential comprehension ability should be adequately developed by fourth grade (Blanchowitz, 1986; Westby, 1984) and therefore the effects of readability and answer format could be examined validly. The children were enrolled in regular education public school classrooms in Seattle, Washington. They all spoke English as their first language and had no auditory impairments or uncorrected visual impairments. In order to insure grade-level reading ability, all participants had scores within the average range (+/- 1.5 standard deviations from the mean) on a standardized, word-reading vocabulary test. Specific subject characteristics are detailed in Table 1. Participating classrooms were those in which the teacher had already agreed to be a part of a larger project designed to examine teacher beliefs about and knowledge of reading instruction and resultant student achievement in reading.

Table 1: Subject Characteristics

Chronological Age		*Ethnicity*		*Gender*		*Word Reading Ability*	
Mean	Range	Race	Percent of sample	Gender	Percent of sample	Gates-MacGinitie Reading Test: Vocabulary Subtest Percentile Rank:	
9.7	8.9-11.1	African American	5	Male	45	Mean	Range
		Hispanic American	13.3	Female	55	65	16-99
		Caucasian	66.7				
		Asian American	11.7				
		Other	3.3				

General Procedures

In order to compile enough reading comprehension questions to support the comparisons made in this study, three different reading tasks were utilized during data collection (the specific nature of the items in each task will be discussed in the following section). The participants completed the multiple-choice Reading Comprehension Subtest of the fourth grade Gates-MacGinitie Reading Test[1] Form L (MacGinitie & MacGinitie, 1989), which contained grade-appropriate reading passages and 45 multiple-choice

[1] These standardized tests along with several other measures (the Vocabulary subtest of the Gates-MacGinitie, the Spelling subtest of the Weschler Individual Achievement Test, an orthographic task, and a written narrative sample) are being administered as part of the teacher beliefs study.

questions. Because short-answer tasks and below-grade-level multiple-choice questions were necessary for the comparisons in this study, the participants also completed two supplementary passage-reading tasks containing twenty-one additional multiple-choice items and fifty-nine short-answer items, respectively. The three tasks were group-administered by this investigator during two sessions in the students' classrooms. All tasks are provided in Appendix A. Two testing assistants were also present to ensure that the participants followed instructions appropriately and completed all items on each task. Task presentation was counterbalanced with respect to answer format and readability level to avoid order effects. All written answers were examined immediately following administration and, if necessary, were verified with students to insure legibility and comprehensibility for later scoring.

Measurements

The three reading tasks administered were comprised of reading passages that varied as to their readability level, and of questions (items) that varied as to readability level (below-grade-level vs. grade-appropriate), answer format (multiple-choice vs. short-answer), and question type (literal vs. inferential). Once the participants completed all three reading tasks, the items were scored and reorganized into eight separate subscales used for the comparisons in this study (See Figure 2)

LITERAL COMPREHENSION	*INFERENTIAL COMPREHENSION*
Readability/Answer Format:	**Readability/Answer Format:**
1.Below-Grade-Level /***Multiple-Choice***	5.Below-Grade-Level /***Multiple-Choice***
2.Grade-Appropriate/***Multiple-Choice***	6.Grade-Appropriate /***Multiple-Choice***
3.Below-Grade-Level/ Short-Answer	7.Below-Grade-Level /Short-Answer
4.Grade-Appropriate/ Short-Answer	8.Grade-Appropriate /Short-Answer

Figure 2: Eight Reading Comprehension Subscales

The eight subscales are described below.

Literal Reading Comprehension Subscales:

Multiple-Choice:

1. Below-Grade-Level Readability: This subscale consisted of 14 items with below-grade-level readability taken from the Reading Comprehension subtests of the third and fourth grade Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1989). The items had readability levels between 1.0 and 3.7. See Appendix B for a discussion of how these readability levels were determined. The items (see example in Table 2) were comprised of

short passages from 3-12 sentences in length followed by questions that required a multiple-choice response. The correct answers for these questions were based upon understanding of the factual material, and the text provided the information necessary to choose the right answer. Specifically, questions were classified as literal "if the student can answer by choosing a restatement of something stated explicitly in the passage (MacGinitie & MacGinitie, 1989, p.10)." Correct answers to these items were summed and the proportion correct was calculated to formulate a "Below Grade Level Readability Literal Score" (# items correct/14).

2. Grade-Appropriate (at or above grade level) Readability: This subscale consisted of 15 items with grade-appropriate readability taken from the Reading Comprehension subtests of the third and fourth grade Gates-MacGinitie Reading Test (Form L) (MacGinitie & MacGinitie, 1989). (MacGinitie & MacGinitie, 1989). These items had readability levels between 4.0 and 6.0, which were determined by the processes discussed in Appendix B. These items (see example in Table 2) were comprised of short passages from 5-15 sentences in length followed by questions that required a multiple-choice response. The correct answers for each item were explicitly given in the reading passage. Correct answers to these items were summed and the proportion correct was calculated to formulate a "Grade-Appropriate Readability Literal Score (# items correct/15)."

Short Answer:

3. Below-Grade-Level Readability

This subscale was comprised of 12 items with below-grade-level readability taken from parallel third and fourth grade forms (Form L) of the Reading Comprehension subtests from the Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1989). The items had readability levels between 1.0 and 3.7 (see Appendix B for how readability was determined). These items (see example in Table 2) were identical to the multiple choice items described immediately above, but the four answer choices that followed the questions were removed. The question stems remained the same and were followed by a series of blank lines that required a short (<1 sentence) written answer. The correct answers to these items were totaled and a proportion correct was calculated to form a "Below Grade Level Readability Literal Score -Short Answer (# items correct/12)."

4. Grade-Appropriate Readability: This subscale consisted of 15 items with grade-appropriate readability taken from parallel forms (Form L) of the Reading Comprehension subtests of the third and fourth grade Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1989). These items had a readability level between 4.0 and 6.0. See Appendix B for a discussion of how these readability levels were determined. These

Table 2: Examples of Multiple-Choice and Short-Answer Literal Comprehension Questions

Literal Multiple Choice Questions	Literal Short-Answer Questions
Below-Grade-Level Readability (Range: 3.0-3.7) Space ships in movies are really just little models of space ships. One movie had a very special camera. It was used to film the little models. It made them look much bigger than they were. And it also made them look as if they were flying. As the camera took pictures it kept moving. Movie people say it was "traveling." Sometimes it went away from a model, backward. Sometimes it went diving under one. On the screen it looked as if the model were flying and climbing. **This story is about space ships in** o museums. o battle. o space. o films.	*Below-Grade-Level Readability (Range 2.3-2.9)* **Ramon made a pencil holder** **He began by painting an empty can. Then he cut out pictures of animals and pasted them on the can.** **What did Ramon do first?** _____ _____
Grade-Appropriate Readability (Range: 5.0-5.9) Mexican American settlers made houses out of what the land had to offer - adobe, or heavy clay. Adobe was molded by mixing the clay with straw and pouring it into molds. The adobe dried in the sun, and it took the shape of a brick. Adobe bricks permitted people to build high thick walls. The Mexicans added windows, fireplaces, and stairs to the homes. The thick adobe brick walls kept out the heat in summer and the cold in winter. **What was added to the adobe?** o Clay. o Straw. o Paint. o Bricks. **What made the soft clay hard?** o A mixer. o The sun. o An oven. o The cold.	*Grade-Appropriate Readability (Range: 5.0-5.9)* *A pueblo child tells how a desert plant is used.* Mothers gather The flowers of guaco, And the leaves and stems; They boil them together On a fire in the plaza; For many days They boil the guaco. They make little, soft cakes Of the boiled-down guaco, And wrap them In corn-husk blankets To dry into paint, Into black paint, To make the designs On pottery. **What is the paint made from?** _____ _____ **What is the paint used on?** _____

reading levels were also determined using the process discussed above. These items (see example in Table 2) were identical to the grade-appropriate multiple-choice items described above, but the answer choices that followed the questions were removed. The question stems remained the same and were followed by a series of blank lines that required a short (<1 sentence) written answer. The correct answers to these items were totaled and the proportion correct was calculated to form a "Grade-Appropriate Readability Inferential Score -Short Answer" (# items correct/15).

Inferential Reading Comprehension Subscales:

Multiple-Choice:

5. Below Grade Level Readability: This subscale consisted of 13 items with below-grade-level readability taken from the Reading Comprehension subtests of the third and fourth grade Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1989). The items had readability levels between 1.0 and 3.7. These reading levels were determined by the procedure discussed in Appendix B. All items (see example in Table 3) were comprised of short passages from 3-12 sentences in length followed by questions that required a multiple-choice response. The correct answers for these 15 items required inferential thinking because "the student can *not* answer the question by choosing a restatement of something stated explicitly in the passage (MacGinitie & MacGinitie, 1989, p. 10)." The correct answers to these items were summed and the proportion correct was calculated to form a "Below Grade Level Readability Inferential Score - Multiple Choice (# items correct/13)."

6. Grade-Appropriate Readability: This subscale consisted of 14 items with grade-level or higher readability taken from the Reading Comprehension subtests of the third and fourth grade Gates-MacGinitie Reading Test (Form K) (MacGinitie & MacGinitie, 1989). These items had readability levels between 4.0 and 6.0 (See Appendix B). All items (see example in Table 3) were comprised of short passages from 5-15 sentences in length that required a multiple-choice response. Again, the correct answers for these 15 items required inferential thinking because the student could *not* answer the question by choosing a restatement of something stated explicitly in the passage. The correct answers to these items were totaled and the proportion correct was calculated to formulate a "Grade Appropriate Readability Inferential Score - Multiple Choice (# items correct/14)."

Short Answer:

7. Below-Grade-Level Readability: This subscale was comprised of 15 items with below-grade-level readability taken from parallel third and fourth grade forms (Form L) of the Reading Comprehension subtests from the Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1989). The items had readability levels between 1.0 and 3.7. These items (see

example in Table 3) were identical to the multiple choice items described immediately above, but the four answer choices that followed the questions were removed. The question stems remained the same and were followed by a series of blank lines that required a short (<1 sentence) written answer. The correct answers to these items were totaled and the proportion correct was calculated to form a "Below Grade Level Readability Inferential Score -Short Answer (# items correct/15)."

Table 3: Examples of Multiple-Choice and Short-Answer Inferential Questions

Inferential Multiple-Choice Questions	Inferential Short-Answer Questions
Below-Grade-Level-Readability (Range: 2.3-2.9) Samuel Duck wanted to build a tree house. "I know just the place for one," he said. "I can build it all by myself, too." Samuel hurried to the woodshed for a ladder. He carried it over to the big apple tree, and up he went. "Oh, do be careful!" called Samantha from below. "Ducks don't belong in trees." **Why did Samuel get a ladder?** o To get to the roof. o To climb a tree. o To go below. o To reach Samantha. **19. Where was Samantha?** o Under the tree. o Up in the tree. o In the woodshed. o Across the street.	*Below-Grade-Level Readability (Range: 2.3-2.9)* It was spring in Peapack. The mayor called a meeting of the town. "Contest at the county fair!" he shouted. "And a cash prize to the town that grows the finest vegetables!" **1. The contest was to see who could** _____ _____ **2. Why did the mayor shout?** _____ _____
Grade-Appropriate Readability (Range: 4.0-4.9) Most glowworms glow with only one color. But the larva of one large beetle has two different sets of colored lights. This is the so-called railroad worm. When it is resting, a pair of bright red lights glows from its head. But if this strange creature begins to crawl along, a row of bright green lights on each side of its body winks on. **The railroad worm is unusual because its lights are** o bright green. o different colors. o in a row. o warm. **A railroad worm is a kind of** o signal light. o tool. o snake. o larva.	*Grade-Appropriate Readability (Range: 4.0-4.9)* Sal was looking through some things out in Gran's shed. Sal dug deep into the trunk. She came up with something she had never noticed before on other visits to Gran's place. A girl's coat. An olden days coat. It was dark navy blue, with a hood, and at the waist there was attached a narrow red wool sash. It looked as though it might fit Sal. She decided to try it on. She slipped out of her own coat and into the old one. It fitted her perfectly. **Where in the trunk was the coat?** _____ **When had the coat probably been worn by Gran?** _____

8. Grade-Appropriate Readability: This subscale consisted of 17 items with grade-appropriate readability taken from parallel forms (Form L) of the Reading Comprehension subtests of the third and fourth grade Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1989). These items had a readability level between 4.0 and 6.0 (See Appendix B). These items (see example in Table 3) were identical to the grade-appropriate multiple-choice items described above, but the answer choices that followed the questions were removed. The question stems remained the same and were followed by a series of blank lines that required a short (<1 sentence) written answer. The correct answers to these items were totaled and the proportion correct was calculated to form a "Grade-Appropriate Readability Inferential Score -Short Answer" (# items correct/17).

Data Reduction

Initially, all items were scored in the context of the original three tasks administered. Once the items were marked as correct or incorrect, they were reorganized into the eight subscales based upon their readability, answer format and question type. Specific guidelines used for scoring are discussed for each answer format:

1. Multiple-Choice:

The multiple-choice items were scored by this author, and correct answers were those dictated by the Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1989). Once the items were scored, they were grouped in the appropriate subscale according to question type and readability level. The correct answers for each subscale were summed and proportions correct were calculated to formulate the following scores:

Below-Grade-Level Readability Literal Score-Multiple Choice (subscale # 1).

Grade-Appropriate Readability Literal Score-Multiple Choice (subscale # 2).

Below-Grade-Level Readability Inferential Score-Multiple-Choice (subscale # 5).

Grade-Appropriate Readability Inferential Score-Multiple-Choice (subscale #6).

See Figure 2 on p. 12 to review subscale numbering and organization.

2. Short-Answer

The correct answers for the short-answer items were based on the multiple-choice answers provided by the Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1989). These answers were expanded to include answers that conveyed the same or a highly similar meaning. For example, if the correct multiple-choice answer to the question "How does silver look when it tarnishes?" was "it gets dark," other acceptable versions of this included "it darkens," "it's a dark color," or "it turns black/brown." See

correct answers in Appendix D. The students' short-answer responses were scored as correct if they matched one of these relevant answer versions.

In order to ensure accuracy in this scoring procedure, the short-answer items were scored by this author, and by one additional individual who was trained in scoring procedures, but who was unaware of the purpose of the study. Once this scorer completed training and demonstrated 100% accuracy in her ability to evaluate a set of trial items correctly, she individually scored the short-answer items. Once this additional scoring was completed, the scoring results were compared for each item. Items for which there was not agreement (9%) were examined, and a collective decision was made as to whether the item should be scored as correct or incorrect. The number of correct items in each subscale was totaled, and proportions correct were calculated to formulate the following scores:

Below-Grade-Level Readability Literal Score – Short Answer (subscale #3).

Grade-Appropriate Literal Score – Short Answer (subscale #4).

Below-Grade-Level Readability Inferential Score - Short-Answer (subscale #7).

Grade-Appropriate Readability Inferential Score - Short Answer (subscale #8).

See Figure 2 on p. 12 to review subscale numbering and organization.

Validity and Reliability

1. Literal Reading Comprehension Subscales:

Multiple-Choice:

Validity: The validity of the literal subscales was addressed prior to their creation in the following ways: First, the Gates-MacGinitie test manual was consulted to determine if the definition of a "literal" question conformed to those found in the research literature. This was indeed the case. The authors of the Gates-MacGinitie Reading Tests created literal or "factual" questions for their Reading Comprehension subtest that require usage of basic information provided in the text. The items were defined as literal if the student "could answer by choosing a restatement of something stated explicitly in the passage." In addition, personal communication with Walter MacGinitie (1998) further confirmed that this definition was theoretically congruent with the definition of literal comprehension utilized in this study (Refer to definitions in Appendix C).

Second, to insure the conformity of each item to this definition, potential literal items were evaluated by two graduate student judges (who have specific experience in language and reading issues) and this author. See the definitions and example items used in this evaluation in Appendix C. Only the items that achieved a three-way consensus as to their literal nature were utilized in the study, providing evidence of construct validity.

Reliability: To determine the internal consistency of both the Below-Grade-Level and the Grade-Appropriate Readability Literal Multiple-Choice subscales, the Split-Half with the Spearman-Brown Prophesy formula (Spearman, 1904) was utilized. This approach involves the creation of two parallel tests by splitting the task items and using the Spearman-Brown Prophesy formula. Reliability coefficients are provided in Table 4, and they were determined to be appropriate for the purposes of this study given that they were all approaching or above .70. This value is an acceptable standard as the reliability coefficient is directly interpretable as percentage of variance without being squared (i.e. 70% of the score variance is due to differences in the performances of the participants). In fact, coefficients as low as .50 are acceptable if tests are being used to make more global decisions about group performance as opposed to specific decisions about individual performance (Worthen, White, Fan, & Sudweeks, 1999).

Short Answer:

Validity: The validity of these subscales was addressed utilizing the procedures discussed above.

Reliability: In order to determine the reliability of both the Below-Grade-Level and the Grade-Appropriate Reading Literal Subscales -Short-Answer, the Split-Half with the Spearman-Brown Prophesy formula (Spearman, 1904) was again utilized and reliability coefficients are detailed in Table 4. While the Grade-Appropriate Short-Answer reliability value is within acceptable limits, the Below-Grade-Level Short-Answer value of .4351 is somewhat low. This value is lower due to the lack of variability in the performance of the students on this easier-to-read subscale. Because many of the students were able to read the material and perform well on the subscale, this truncated range created a lower reliability coefficient.

Table 4: Split-Half with the Spearman Brown Reliability Coefficients for All Literal Subscales

Literal Subscale	***Reliability Coefficient***
Below-Grade-Level Multiple-Choice	.6719
Grade-Appropriate Multiple-Choice	.7208
Below-Grade-Level Short-Answer	.4351
Grade-Appropriate Short-Answer	.7085

2. Inferential Reading Comprehension Subscales:

Multiple-Choice:

Validity: The validity of these items was also addressed prior to the creation of the subscales. Again, the Gates-MacGinitie test manual was consulted to examine the definition of inferential questions, and the process discussed above was followed. The authors of the Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1989) created

inferential items for their Reading Comprehension subtest that require application of knowledge beyond that provided in the text. Specifically, inferential questions are defined as "those questions that could not be answered by choosing a restatement of something stated explicitly in the passage." Personal communication with Walter MacGinitie (1998) further confirmed that this definition was theoretically congruent with the definition of inferential comprehension utilized in this study. The inferential nature of the questions from the Gates-MacGinitie Reading Comprehension subtest were confirmed by two graduate student judges who have specific experience in language and reading issues and by this author. After specific instruction in inferential issues and the definition of inferential items utilized in this study (see Appendix C), the judges rated all potential items. Only items that were unanimously judged to be inferential were included in the study.

Reliability: In order to determine the reliability of both the Below-Grade-Level and the Grade-Appropriate Reading Inferential subscales – Multiple-Choice, the Split-Half with the Spearman Brown Formula (Spearman, 1904) was utilized. The reliability coefficients are detailed in Table 5 and both values were approaching or above .70 and were deemed appropriate for use of the tasks in the study.

Short Answer:

Validity: The validity of these subscales was addressed utilizing the procedures discussed above.

Reliability: In order to determine the reliability of both the Below-Grade-Level and the Grade-Appropriate Reading Inferential Tasks -Short-Answer, the Split-Half with the Spearman-Brown Prophesy formula (Spearman, 1904) was utilized. Reliability coefficients are provided in Table 5. While the Grade-Appropriate Short-Answer reliability value is within acceptable limits, the Below-Grade-Level value of .3839 is somewhat low. This value is lower due to the lack of variability in the performance of the students on this easier-to-read subscale. Because many of the students were able to read the material and perform well, this truncated range created a lower reliability coefficient.

Table 5: Split-Half with the Spearman Brown Reliability Coefficients for all Inferential Subscales

Inferential Subscale	***Reliability Coefficient***
Below-Grade-Level Multiple-Choice	.6224
Grade-Appropriate Multiple-Choice	.7746
Below-Grade-Level Short-Answer	.3839
Grade-Appropriate Short-Answer	.7449

In addition to the above measure of internal consistency, *Inter-Rater Reliability* was calculated to insure that item scoring was consistent for the short-answer questions. The

scores generated by the independent scorer in the procedure discussed above (see p. 18) were compared to scores calculated by this author. Correlations between these scores are found in Table 6. All reliability coefficients were above .90 and significant at the 0.01 level.

Table 6: Inter-rater Reliability Values for Short-Answer Tasks

Short-Answer Subscale:	***Correlation Between Both Scorers:***
Below-Grade-Level Literal	.978
Below-Grade-Level Inferential	.978
Grade-Appropriate Literal	.976
Grade-Appropriate Inferential	.990

Item Fit Analysis

The fit of each item to both a "total test" score (on all 115 items) and to its individual subscale score were examined. The former determines if the items within a "total test" work together to measure a general construct (i.e. reading comprehension performance), and then as a subscale to measure a more specific construct (i.e. reading comprehension performance on below-grade-level literal multiple-choice items). Biserial item-test correlations were used to correct for dichotomous data. To examine the fit of each item with a "total test" score, the biserial correlation of the correct responses on each item to a "total test" score was calculated. The mean item-"total test" correlation was 0.48, indicating that the items within "total test" demonstrated marginal fit. This value may have been reduced due to the lack of variability of subjects' performances on some of the below-grade-level items. To examine the fit of each item within its particular subscale, biserial correlations were calculated between correct responses to each item and the number correct scores on each subscale to which the item was assigned. The mean biserial values (Table 7) for each subscale were stronger, ranging from 0.514 to 0.681, indicating that the items were more homogeneous within their subscales than within the "total test," and providing evidence for their discriminant validity. Some of these values may also have been reduced due to the lack of variability of subjects' performances on some of the below-grade-level items.

Table 7: Mean Biserial Correlations Between Item and Subscale

Subscale:	***Mean Biserial Correlation***
1. Below-Grade-Level Multiple-Choice Literal	.681
2. Grade-Appropriate Multiple-Choice Literal	.622
3. Below-Grade-Level Short-Answer Literal	.514
4. Grade-Appropriate Short-Answer Literal	.640
5. Below-Grade-Level Multiple-Choice Inferential	.613
6. Grade-Appropriate Multiple-Choice Inferential	.633
7. Below-Grade-Level Short-Answer Inferential	.531
8. Grade-Appropriate Short-Answer Inferential	.545

Chapter IV

Results

The purpose of this study was to examine the effects of text readability level and answer format on fourth-grade readers' ability to demonstrate comprehension of literal and inferential meaning during reading. A 3-factor repeated measures analysis of variance (ANOVA) was utilized to examine effects of readability (below-grade-level vs. grade-appropriate), answer format (multiple-choice vs. short-answer) and question type (literal vs. inferential) on the reading comprehension performance of 60 fourth graders. Interactions among these variables were also examined. Mean proportion correct scores and standard deviations are listed in Table 8 and mean proportion correct scores are displayed graphically in Figure 3.

There was a significant main effect for question type [$F(1,59) = 139.18$; $p < .001$] with a moderate effect size (.58). [2] Readability had a significant main effect [$F(1,59) = 84.174$; $p < .001$] with a moderate to large effect size (.70). There was a significant main effect for answer format [$F(1,59) = 4.126$; $p < .05$] with a small effect size (.07). All two-way interactions were also significant at the .01 level.

Three-Way Interaction of Readability by Answer Format by Question Type

These main effects and two-way interactions must be qualified in the context of a significant three-way interaction of Readability by Answer Format by Question Type [$F(1,59) = 6.275$; $p < .05$]. The three-way interaction (see Figure 3) indicated differential effects of answer format on literal and inferential comprehension performance at different readability levels. The graph of the interaction in Figure 3 suggests that, at the below-grade readability level, participants' performance on literal questions was better than on inferential questions, and that performances in the multiple-choice format were better than in the short-answer format for both question types. At the grade-appropriate readability level; however, performance on literal questions was better in the short answer format than in the multiple-choice format. The effect size for this interaction was considered small (.10). In light of the interaction and its small effect size, specific mean comparisons were conducted to determine where significant differences were present. Specifically, the effects of answer format and question type were examined separately for both below-grade-level and grade-appropriate readability levels in eight one-tailed Bonferroni post hoc comparisons (Critical t value: 2.89). Differences between subscales with different readability levels were not examined, as the effect of readability in the repeated measures ANOVA was considered moderate to large (.70).

Comparisons between the *answer formats* revealed significant differences between *below-grade-level literal* multiple-choice (Mean 1 in Table 8) and short-answer (Mean 3 in Table 8) subscales [$t(59) = 3.4$], and between *grade-appropriate literal* multiple-choice (Mean 2 in Table 8) and short-answer (Mean 4 in

---

[2] Effect size is considered small at a value of .20, medium at a value of .50 and large at a value of > .80. Small and medium effect sizes are very common in social science research (Stevens, 1996).

Table 8) subscales [t(59) = -3.09]. In the former comparison, participants performed significantly better on the multiple-choice questions, and in the latter comparison, participants performed significantly better on the short-answer questions. No significant differences between answer formats were found for below-grade-level and grade-appropriate *inferential* subscales.

Further comparisons examined differences between *question types* and revealed significant differences between *below-grade-level multiple-choice* literal and inferential subscale scores [t(59) = 6.24], between *below-grade-level short-answer* literal and inferential subscale scores [t(59) = 10.47], and between *grade-appropriate short-answer* subscales [t(59) = 7.67]. The mean difference between question types on the grade-appropriate multiple-choice subscales was not significant.

Table 8: Group Mean Proportion Correct Scores and Standard Deviations for All Subscales

Subscale	*Mean Proportion Correct Score*	*Standard Deviation*
1. Below-Grade-Level Multiple-Choice Literal	88.93	12.30
2. Grade-Appropriate Multiple-Choice Literal	68.22	19.58
3. Below-Grade-Level Short-Answer Literal	84.58	11.67
4. Grade-Appropriate Short-Answer Literal	75.11	19.48
5. Below-Grade-Level Multiple-Choice Inferential	79.36	17.32
6. Grade-Appropriate Multiple-Choice Inferential	67.02	21.85
7. Below-Grade-Level Short-Answer Inferential	70.33	13.91
8. Grade-Appropriate Short-Answer Inferential	62.16	20.04

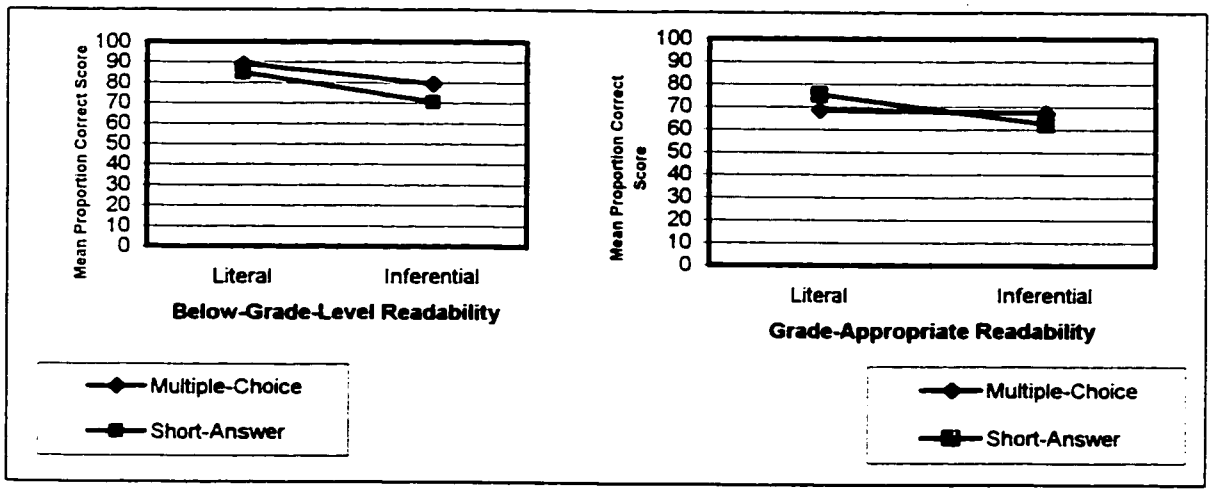


Figure 3: Three-Way Interaction of Readability by Answer Format by Question Type

Item Effects of Readability, Answer Format, and Question Type

To examine the effects of question type, answer format, and readability level in the context of the subscale items, a 2 X 2 X 2 factorial ANOVA was conducted (means and standard deviations are presented in Table 9). There were significant main effects for readability [$F(1,114) = 21.71$; $p < .00$], with below-grade-level items having higher scores; and for question type [$F(1,114) = 12.93$; $p < .00$], with literal items having higher scores. Effect sizes for these two factors were .17 and .11, respectively. Answer format was not significant ($p < .238$) and there were no significant interactions.

Table 9: Means and Standard Deviations for Item Scores

Item Type	*Mean Total Correct*	*Standard Deviation*
1. Below-Grade-Level Multiple-Choice Literal	53.50	5.00
2. Grade-Appropriate Multiple-Choice Literal	41.06	10.18
3. Below-Grade-Level Short-Answer Literal	50.58	8.06
4. Grade-Appropriate Short-Answer Literal	44.93	6.10
5. Below-Grade-Level Multiple-Choice Inferential	47.46	7.85
6. Grade-Appropriate Multiple-Choice Inferential	40.21	6.29
7. Below-Grade-Level Short-Answer Inferential	42.00	14.21
8. Grade-Appropriate Short-Answer Inferential	37.00	7.80

Chapter V

Discussion

Comprehensive assessment is a necessary component of effective reading instruction. Since the goal of instruction is to enable students to comprehend a variety of texts independently, one of the most important questions in comprehension assessment is "How well are they achieving this goal?" The answer to this question leads to a further question, i.e. "How can teachers help students comprehend better? "(Maria, 1990). If the ultimate goal of assessment is to inform and guide this instructional decision-making, the process must encompass a variety of measurement tools that provide accurate insights into both literal and inferential comprehension processing. One commonly used measurement context is characterized by the question-answering tasks (i.e. read a passage and answer questions that follow) often found in standardized reading tests or basal reader evaluations. Very often, these types of tasks do not evaluate literal and inferential comprehension separately, but instead provide a "total" comprehension score that has less instructional value. Teachers need to be aware of the *nature* of students' comprehension strengths and difficulties so that reading instruction can be tailored accordingly.

In addition, there are many factors, both intrinsic and extrinsic to the reader that may influence the effectiveness of this assessment approach. Intrinsic factors include motivation, semantic knowledge, syntactic knowledge, and decoding ability, and extrinsic factors include text characteristics and task requirements. This study was concerned with two *extrinsic* factors, text readability and answer format, that can affect literal and inferential comprehension performance in a question-answering context.

Specifically, this study examined the effects of *readability* and *answer format* on fourth graders' literal and inferential reading comprehension performance. Inferential comprehension was of particular interest because of its critical contribution to reading for meaning. There was a preliminary objective: to confirm differences between literal and inferential comprehension so that examination of them as separate skills would be a valid endeavor. Then, there were two major objectives: to investigate effects of readability on inferential comprehension performance (i.e. Does readability affect inferential comprehension differently than literal comprehension and, if so, how?), and to investigate the role of answer format on inferential reading comprehension performance (i.e. Does answer format differentially affect literal and inferential comprehension performance and, if so, how?).

Literal and Inferential Comprehension Are Different

The study initially compared literal and inferential comprehension to verify presumed differences between them, and to determine if such differences were identifiable in the context of commonly used classroom comprehension activities (i.e. group-administered, written tasks in which students read passages and answer questions). As anticipated, there was a significant difference between literal and inferential comprehension: the item ANOVA provided evidence that inferential items were more difficult and the post hoc comparisons within the repeated measures ANOVA three-way interaction supported that

participants had significantly greater difficulty answering inferential questions in all but the grade-appropriate multiple-choice format. These results are consonant with previous research (Hansen, 1981; Holmes, 1984) indicating that inferential comprehension is more difficult than literal comprehension.

In addition, these results also provide evidence that the difference between literal and inferential comprehension is identifiable in more typical classroom reading activities (as opposed to contrived one-on-one research contexts), so that teachers can evaluate them both utilizing everyday reading materials or commercially available reading achievement tests.

Implications for Assessment and Instruction: With this additional evidence that the two types of comprehension are unique, the implications for assessment are clear. Reading evaluations should consider literal and inferential comprehension separately, because they appear to be different skills that will lend themselves to different instruction. Knowing that students are able to gain literal meaning from a passage provides only partial insight into their abilities as readers. The literal information allows them to form basic propositions (idea units), and perhaps begin to link these propositions together to generate the surface meaning of a passage. However, students will not become successful readers based on this skill alone. The key to reading for meaning is the ability to "read between the lines" and create a situation model (Kintsch, 1978) that characterizes a thorough understanding of a passage or text. Teachers must be able to determine whether or not students are developing this higher-level inferential comprehension skill.

Given the differences between literal and inferential comprehension, students will need exposure to and instruction in how to answer both types of questions. Strategies such as looking for the "who, what, when and where" information, or rereading text to find specific facts will guide literal comprehension but will not help students infer. Strategies such as teaching students how to make predictions during reading, or illustrating how individual words add meaning to a sentence and can be used to activate knowledge will foster inferential thinking during reading. In addition, utilizing a "puzzle metaphor" that prompts metacognitive skills can help students "put the pieces together" through self-talk (e.g. Is there a sentence in the book that gives the answer? If "No" then this is a puzzle question. What information in the book will help me? What else do I know about _____ that will help me? How does this information fit together to make an answer? Does my answer make sense?) (Milosky & Ford, 1997).

Readability Affects Comprehension

Given that literal and inferential comprehension appear to be different, the focus now turns to variables that can affect the measurement of these skills. The level of text difficulty, or readability, appears to affect both literal and inferential comprehension. In the present study, text readability was measured at two levels, below-grade-level and grade-appropriate. As expected, readability of the material had a significant effect both in the repeated measures ANOVA analysis and in the item ANOVA analysis. The fact that grade-appropriate readability levels are more challenging lends support to the theory behind working memory and the LaBerge and Samuels model of reading (LaBerge & Samuels, 1974). Rapid, automatic

decoding of words is necessary so attentional and cognitive resources can be focused on comprehension. Reading more difficult, grade-appropriate material appears to deplete these resources at a more basic level in the reading process. When greater effort is expended for decoding, higher-level comprehension processes (i.e. linking propositions, generating inferences) are less efficient, making successful comprehension more challenging.

Implications for Assessment/Instruction: With respect to accurate assessment of literal and inferential comprehension, readability must be considered so that decoding ability isn't a confounding variable. A variety of readability levels (i.e. beginning at a below-grade-level and graduating to grade-appropriate reading material) should be utilized to determine the reading level where comprehension becomes compromised (due to a lack of automaticity in decoding). Then, instruction can be tailored so reading skills are taught and practiced at a manageable reading level (i.e. the *independent* readability level, at which a student can read successfully without assistance). At this level, students will be able to *decode* well enough to benefit from *comprehension* instruction. While readability is not the only influential variable in the comprehension process, its impact on classroom reading assessment and ultimately on comprehension instruction should not be ignored.

Answer Format Has Different Effects on Comprehension Performance

Although the effects of readability on reading comprehension are straightforward (grade-appropriate readability is always more difficult) and significant, the effects of another variable, answer format, are more complicated and less powerful. While the item ANOVA analysis did not reveal a significant main effect for answer format, post hoc comparisons in the context of the repeated measures three-way interaction revealed two significant mean differences that are worth consideration. Comparisons between the answer formats revealed significant differences between *below-grade-level literal* short-answer and multiple-choice subscales, and between *grade-appropriate literal* multiple-choice and short-answer subscales. Figure 4 shows answer formats for which participants demonstrated better performance (asterisks indicate significant mean answer format differences). In the context of this study, literal questions at both readability levels showed significant answer format differences, but inferential questions did not.

	Readability Level	
Question Type	Below-Grade-Level	Grade-Appropriate
Literal	Multiple-Choice *	Short-Answer *
Inferential	Multiple-Choice	Multiple-Choice

Figure 4: Answer Formats at Which Participants Demonstrated Better Comprehension Performance

Performance in the short-answer format may have been better for literal questions at a grade-appropriate reading level because, with the greater decoding challenge, participants relied more heavily on their ability to successfully gain literal meaning as they read. Perhaps the short-answer format then allowed them to write literal answers quickly and from memory. In addition, there was less risk of participants second-guessing themselves due to multiple-choice distracter items (i.e. the choice that is plausible but not exactly accurate), or due to the additional time and processing required for reading multiple answer choices. Another possibility is that, with the increased semantic and syntactic complexity of grade-appropriate reading material, participants mentally represented literal information in a particular form as they read, but were unable to find an answer that matched this representation in the multiple choices (i.e. they remembered from the passage that "as the creature begins to *crawl along*, green lights come on," but for the question "Green lights come on when the railroad worm _____?" the correct multiple-choice answer was worded as "*moves*") . These risks may not have been problematic at the *below-grade reading level* because less effort was required to process both the passages and the questions. Instead, the multiple choices may have been simple and straightforward enough to serve as a helpful confirmation of the literal knowledge participants had generated during reading.

While these differences in performance were significant for literal questions, they were not for inferential questions at either grade level. Perhaps answer format differences were not significant for inferential subscales because the greater difficulty of these questions precluded better performance on one answer format over the other. If this were the case, neither the presence of the correct answer within the context of four multiple-choices nor the opportunity to write an answer quickly and from memory appeared helpful enough to significantly improve performance. Participants may have been able to select a multiple-choice answer or write in a short-answer equally well given the challenge of answering the inferential questions.

Implications for Assessment/Instruction: Overall, answer format had the least significant effects on comprehension performance, yet the idea that students' performance may be affected even in the slightest due this variable makes it worth further consideration with respect to comprehension assessment and treatment.

Conclusions

In summary, reading comprehension assessment tools range from the unstructured, spontaneous gathering of information during instruction, to structured tests with specifically defined outcomes and directions for administration and scoring (e.g. standardized reading comprehension tests) (Maria, 1990). Each method should be considered with respect to the information it is capable of providing. This study examined factors that should be taken into consideration when utilizing a more structured, question-answering paradigm. To be thorough, this assessment should involve examination of both literal and

inferential comprehension and, to be valid, this assessment should be conducted with the knowledge that readability level and answer format affect student performance. If teachers don't have a true measure of comprehension ability, instructional efforts may not be as effective.

Limitations

In interpreting the research findings, several limitations to the present study should be noted. First, the subscales had a limited number of items, and may not have been comprehensive enough for accurate measurement of each type of comprehension at each readability level and in each answer format. There was a limited pool of items to choose from in the Gates-MacGinitie Reading Comprehension subtests, and several of theses items were removed when they did not conform to the definitions of literal and inferential comprehension used in the study. In addition, classroom time limits and participant energy constraints had to be considered, as participants were asked to read approximately 50 passages and answer approximately 130 questions. While there is no readymade formula that dictates how many items are necessary to obtain a representative sampling of a behavior, the small number of items representing each specific reading skill may have limited accurate analysis of literal and inferential comprehension abilities. In addition, the subscales were not standardized instruments, thus limiting the application and generalization of participant performance.

A second limitation was that the two below-grade-level short-answer subscales had limited split-half reliability, potentially influencing their use in the analysis. Because of the limited number of subscale items, and because many of the students performed well on these subscales and truncated the range, a lower reliability coefficient resulted. The analyses involving scores on these subscales should be interpreted with caution.

Directions for Future Research

The present study provides evidence that readability and answer format should be taken into account when reading comprehension is assessed. Given these conclusions, future studies could address issues regarding readability and answer format, and issues related to reading comprehension in general. One possible avenue of research would be to further examine the use of a multiple-choice vs. a short-answer format, and try to uncover why, in the present study, performance on these formats did not differ when students were answering inferential questions, but did differ when they were answering literal questions. Answers to inferential questions could be compared to student thinking throughout the reading process. Participants could be engaged in "think aloud" procedures (Suh & Trabasso, 1993) during which they talk out loud about their comprehension as they read. They could be asked to continue this process as they answer both multiple-choice and short-answer inferential comprehension questions following the passage. This type of information would provide insight into inferential thinking during reading (i.e. Are inferences made on-line?), how it relates to students' ability to answer comprehension questions (Are

students strategic in their approach to answering both multiple-choice and short-answer inferential questions?), and ways to better assess inferential ability in reading comprehension evaluations.

In addition, studies might take into account the potential limitations of having students generate *written* answers in a short-answer format. A comparison of oral and written answers to inferential questions would sort out differences in inferential reasoning ability versus difficulties conveying ideas within the spelling, handwriting, or organizational constraints of written language. The effects of both multiple-choice and short-answer formats on performance of students with language/learning or reading disabilities could also be examined. This would provide additional insight into the impact reading and language abilities have on inferential reading comprehension performance, and may also identify different issues in assessment of and instruction for children with a variety of academic needs.

Another means of investigating answer format would involve examination of the nature of student errors when generating short answers to both literal and inferential questions. Oral or written answers could be analyzed qualitatively to examine mistakes and determine the rationale behind them, providing additional insight into potential breakdowns in literal and inferential comprehension short-answer performance.

A second avenue of research might take the information from this study and apply it in the context of inferential comprehension instruction. When readability and perhaps even answer format are manipulated to provide for optimal student learning, the growth in comprehension ability could be monitored and compared to students' performance when these factors were not take into account.

A third avenue of research might look more deeply into the process of inferential reading comprehension and, specifically, what elements in this process make it different from literal comprehension. The extrinsic factors investigated in this study (readability and answer format) had effects on both literal and inferential comprehension performance, so there must be additional variables that make the latter more challenging. Factors that are intrinsic to the reader such syntactic knowledge, semantic knowledge, or world knowledge could be examined to determine their contribution to both literal and inferential reading comprehension performance. Perhaps these skills are relied upon more heavily when readers attempt to generate inferences than when they try to attain literal meaning. This type of investigation could be completed with both normally-achieving students and individuals with learning disabilities to gain insight into what skills are necessary for and what happens when there are breakdowns in inferential reading comprehension.

References:

Blanchowitz, C. (1986). Semantic constructivity in children's comprehension. *Reading Research Quarterly*, 13, 188-199.

Casteel, M. (1993). Effects of inference necessity and reading goal on children's inferential generation. *Developmental Psychology, 29*, 2, 36-357.

Chronback, L.J. Coefficient alpha and the internal structure of tests. *Psychometrika*, 1951, 16, 297-334.

Dale, E. & Chall, J. S. (1968). A formula for predicting readability. *Educational Research Bulletin, 27*, 11-20, 37-54.

Fry, E. B., Fountoukidis, D. L., & Polk, J. K. (1985). *The new reading teacher's book of lists* (pp. 273-274). Englewood Cliffs, NJ: Prentice-Hall.

Hansen, C. (1981). The effects of inference training and practice on young children's comprehension. *Reading Research Quarterly, 16*, 391-417.

Harris, A.J. & Jacobson, M. D. (1982). *Basic reading vocabularies.* New York: Macmillan.

Holmes, B. (1984). The effect of prior knowledge on the question answering of good and poor readers. *Journal of Reading Behavior. 15*, 1-18.

Just, M.A. & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329-354.

LaBerge, D. & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323.

Leslie, L. & Caldwell, J. (1995). *Qualitative Reading Inventory - II.* HarperCollins College Publishers: New York.

Lipson, M. & Wixon, K. (1997). *Assessment and Instruction of Reading and Writing Disability* Longman: New York.

MacGinitie, W. & MacGinitie, R. (1989). *The Gates MacGinitie Reading Tests.* Riverside Publishing Company: Maryland.

Maria, K. (1990). *Reading comprehension instruction: Issues & Strategies.* York Press: Maryland.

Martinez, M. (1991). A comparison of multiple-choice and constructed figural response items. *Journal of Educational Measurement, 28*, 2, 131-145.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99*, 440-466.

Milosky, L. & Ford, J. (1996). *Comprehending what is not said: Critical inferencing skills in comprehension.* Presentation at the American Speech and Hearing Association Convention. Seattle.

Mitchell, D. C. (1982). *A cognitive analysis of fluent reading and learning to read.* New York: John Wiley & Sons.

Murray, J., Klin, C., & Myers, J. (1993). Forward inferences in narrative text. *Journal of Memory and Language, 32*, 464-473.

Oakhill, J. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology, 54*, 31-39.

Omanson, R. C., Warren, W. H., & Trabasso T. (1978). Goals, inferential comprehension, and recall of stories by children. *Discourse Process, 1*, 337-354.

Paris, S. G., & Carter, A. (1973). Semantic and constructive aspects of sentence memory in children.

Development Psychology, 9, 109-113.

Paris, S. G., Lindauer, B. K., & Cox, G. (1977). The development of inferential comprehension. *Child Development, 48,* 1728-1733.

Reder, L. M. (1980). The role of elaboration in the comprehension and retention of prose: A critical review. *Review of Educational Research, 50,* 5-53.

Samuels, S. J. (1987). Information processing abilities and reading. *Journal of Learning Disabilities, 20,* 1, 18-22.

Sanford, A.J., & Garrod, S.C. (1981). *Understanding Written Language: Explorations in comprehension beyond the sentence.* New York: Wiley.

Schachter, S. H. (1980). *An investigation of the effects of vocabulary instruction and schemata orientation upon reading comprehension.* Paper presented at the International Reading Association Conference, St. Louis, Missouri.

Schreiner, R., & Shannon, P. (1980). *The recall of explicit and implied propositions by good and poor readers using three types of assessment procedures.* Paper presented at the annual convention of the International Reading Association, St. Louis, Missouri.

Snow, R. E. (1980). Aptitude processes. In R. E. Snow, P. A. Federico, & W. E. Montague (Eds.), *Aptitude, learning, and instruction, Volume 1: Cognitive process analyses of aptitude* (pp. 27-63). Hillsdale, NJ: Lawrence Erlbaum Associates.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15,* 72-101.

Stathmann, S. (1977). Testing reading comprehension: Multiple-choice v. short-answer questions. *English Language Teaching Journal, 33,* 4, 304-306.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences.* Mahwah, NJ: Lawrence Erlbaum Associates

Suh, S. & Trabasso, T. (1993). Inferences during reading: converging evidence form discourse analysis, talk-aloud protocols, and recognition priming. *Journal of Memory and Language, 32,* 279-300.

Tierney, R. J., Bridge, C., and Cera, M. J. (1978). The discourse processing operations of children. *Reading Research Quarterly,* 4, 541-569.

Thorndike, E. (1917). Reprinted in 1973 by Thorndike, R. L. Reading as reasoning. *Reading Research Quarterly, 9,* 135-147.

Traub, R. E., & MacRury, K. (1991). Multiple-choice vs. free-response in the testing of scholastic achievement. In K. Ingendamp & R. S. Gager (Eds.), *Tests and trends 8: Jahrbuch der padagogischen diagnostik.* Weinheim und Basel, Germany: Beltz Verlag.

Westby, C. (1984). Development of narrative language abilities. In G. P. Wallach & K. Butler (Eds.), *Language learning disabilities in school age children.* (pp. 103-127). Baltimore: Williams and Wilkins.

Wilson, M. M. (1979). The processing strategies of average and below average readers answering factual and inferential questions on three equivalent passages. *Journal of Reading Behavior, 11,* 235-245.

Worthen, B., White, K, Fan, X., Sudweeks, R. (1999). *Measurement and Assessment in Schools.* New York: Longman.

Zabrusky, K. (1986). The role of factual coherence in discourse comprehension. *Discourse Processes, 9,* 197-220.

Appendix A: Literal and Inferential Multiple-Choice and Literal and Inferential Short-Answer Tasks

TASK 1: Comprehension Subtest of the Fourth-Grade Gates MacGinitie Reading Test (Form L)
(Circled items are inferential)

Comprehension

Eohippus (Ee-o-*hip*-us) was an animal that lived long long ago.

It was not much bigger than a rabbit.

It lived in the open. It ate plants that grew in the open. It ran very fast. It had to run fast, because there were many bigger animals that wanted to eat it.

Eohippus was the first horse in the world. Even though it was very small, it was really a horse. All of our horses today come from Eohippus.

1. The first horses were about as big as one of today's

   ⓐ horses. ⓒ mice.
   ○ rabbits. ⓓ Eohippus.

2. Where did Eohippus live?

   ⓐ In jungles.
   ⓑ In the sea.
   ⓒ In holes in the ground.
   ○ In open places.

3. Eohippus needed to go fast so it could

   ⓐ be first.
   ⓑ get out of danger.
   ○ catch its food.
   ⓓ stay with bigger animals.

Prince Roderick was visiting his cousin Krispin. But Krispin lived out in the country. And Roderick was used to a fine city palace.

For supper Krispin served a good, hot meal.

But Roderick only picked at his food.

"You'll never guess what's going to happen tomorrow," Krispin said. "It's a fair. We're all going to have fun! There will be good food, singing, and turtle racing. You do like turtle racing, don't you?"

"No!" Roderick growled.

4. Roderick thought the meal was

   ⓐ not what he wanted.
   ⓑ too small.
   ⓒ too hot.
   ⓓ late.

5. Roderick growled because he was

   ⓐ pretending to be a lion.
   ⓑ having fun.
   ⓒ being unpleasant.
   ⓓ hungry.

GO ON

It was a good thing for Neesie that Mama couldn't see her, or she would have got a good spanking.

Mama couldn't hear her either, but I could. All the time Mama was cornrowing my hair, Neesie kept calling me and waving her arms around, trying to make me look at her. After a while, I got tired of it.

"Stop it, Neesie!" I said.

6. How was Neesie acting?

Ⓐ Tired. Ⓒ Naughty.
Ⓑ Polite. Ⓓ Surprised.

7. Neesie wanted the girl who is telling the story to

Ⓐ get a spanking.
Ⓑ sing.
Ⓒ dance.
Ⓓ watch.

8. The girl telling the story got tired of

Ⓐ listening to Neesie.
Ⓑ playing with Neesie.
Ⓒ calling.
Ⓓ waiting.

9. What was Mama doing?

Ⓐ Waving her arms.
Ⓑ Yelling at Neesie.
Ⓒ Fixing the girl's hair.
Ⓓ Getting dinner ready.

Most glowworms glow with only one color. But the larva of one large beetle has two different sets of colored lights. This is the so-called *railroad worm*. When it is resting, a pair of bright red lights glows from its head. But if this strange creature begins to crawl along, a row of bright green lights on each side of its body winks on.

10. The railroad worm is unusual because its lights are

Ⓐ bright green.
Ⓑ different colors.
Ⓒ in a row.
Ⓓ warm.

11. A railroad worm is a kind of

Ⓐ signal light.
Ⓑ tool.
Ⓒ snake.
Ⓓ larva.

12. Green lights come on when the railroad worm

Ⓐ moves. Ⓒ stops.
Ⓑ rests. Ⓓ swims.

13. This worm is called a *railroad* worm because of

Ⓐ the sounds it makes.
Ⓑ what it looks like.
Ⓒ the tracks it goes on.
Ⓓ the way it pulls things.

GO ON

My mother is like the sun rising
in the early morning,
lighting up the dark corners
and gently coaxing us awake.
She prods the fire into life
and soon everywhere is filled
with the smoky smell of food,
bringing rumbles to my tummy
and making me want to get up.

14. The child thinks his mother is like the sun because of the way she

(M) is everywhere.
(N) sings to him.
(O) wakes him up.
(P) keeps him warm.

15. The mother prods the fire to

(Q) move it to a different place.
(R) keep it under control.
(S) make it smoky.
(T) get it started.

16. Why does the child want to get up?

(A) Everyone else is up.
(B) It is too smoky.
(C) He is hungry.
(D) He is in a dark corner.

When I was a little girl, I wanted to become a doctor, and a dentist is one type of doctor. I decided to become a children's dentist because I enjoy working with my hands and I like children and want to help them. There are many parts of my job that I like. One of the nicest is seeing a patient who at first didn't want to visit the dentist come back again with a smile.

17. The writer is a dentist partly because

(E) she can work with her hands.
(F) she was first a doctor.
(G) she is little.
(H) her father was a dentist.

18. Who comes back again with a smile?

(I) A doctor.
(J) A dentist.
(K) A child.
(L) A friend.

19. The person comes back again with a smile because the person

(M) has white teeth.
(N) likes to come.
(O) is nicest.
(P) is first.

GO ON

Paul had been at Frobisher School for nearly three years. He was now twelve. He had become so much a part of the place that he spoke English all the time and thought and even dreamed in English. He liked it there, for he had learned how to play games. He played football in the fall and hockey from December to March. In the early spring there was boxing, and the coach said that he was a natural at it.

20. When did Paul start going to Frobisher School?

- (G) In the early spring.
- (H) When he was twelve.
- (J) Before his friends went.
- (K) Before he was twelve.

21. What did Paul do during the winter?

- (A) He went home.
- (B) He went skiing.
- (C) He practiced football.
- (D) He played hockey.

22. The coach helped Paul by

- (E) encouraging him.
- (F) warning him.
- (G) listening to him.
- (H) sending him to Frobisher School.

23. What is this passage mainly about?

- (I) How old Paul was.
- (J) How well Paul spoke.
- (K) How Paul felt about school.
- (L) What Paul hoped to do.

Silk comes from cocoons silkworms spin. Silkworms are the caterpillars of certain moths. Silkworms feed on mulberry leaves. Then they spin cocoons. When they come out of the cocoons they are full-grown moths.

Many cocoons are gathered before the moths come out. These cocoons are cooked until they are soft. Their fine soft thread is unwound. Several of the fine threads are spun together to make stronger ones. Then many of them are woven into silk cloth.

24. These caterpillars are called silkworms because they

- (F) lay eggs on silk.
- (G) make silk.
- (H) look like silk.
- (J) feed on silk.

25. The cocoons are cooked to make them

- (A) clean.
- (B) stronger.
- (C) soft.
- (D) good to eat.

26. When do people gather the cocoons?

- (A) Before the silkworms feed on the leaves.
- (B) When the silkworms are in the cocoons.
- (C) When the threads are fine enough.
- (D) When the silk comes from the cocoons.

27. When is the silk unwound?

- (E) Before the cocoon is spun.
- (F) After the moth comes out.
- (G) After the threads are made stronger.
- (H) After the cocoon is cooked.

GO ON

Comprehension

Janie was in a motorbike race.

The bikes in front of Janie disappeared over the top of the hill. Janie braked her bike at the top, then with a rush she charged down the steep hill.

Dirt and rocks flew from under the bike's wheels. From the corner of her eye, Janie saw that one bike had fallen. The rider was tumbling down the hill. Janie braked her bike again.

Soon Janie's bike started to slide. She let up on the brake and turned up the gas. Her bike shot ahead.

28. The other bikes disappeared because they

- ⓘ were stolen.
- ⓙ were far behind.
- ⓚ were behind the hill.
- ⓛ had fallen.

29. Rocks flew because

- ⓜ the bike had fallen.
- ⓝ the bike was going fast.
- ⓞ the people were angry.
- ⓟ Janie braked her bike.

30. To stop her bike from sliding, Janie

- ⓠ speeded up.
- ⓡ leaned forward.
- ⓢ braked it.
- ⓣ put her feet on the ground.

31. Janie's bike shot ahead because

- Ⓐ the hill was too steep.
- Ⓑ the bike started to slide.
- Ⓒ a rider was tumbling down the hill.
- Ⓓ Janie gave it more gas.

Dawn is coming to the city. Slowly the eastern sky begins to brighten. Here a light goes on . . . there a light goes on . . . . The city sparrows begin to cheep. The ducks on the pond in the park call to each other across the black water.

32. Why do the lights go on?

- Ⓔ It is getting darker.
- Ⓕ The sparrows begin to cheep.
- Ⓖ People are afraid.
- Ⓗ People are getting up.

33. What are the ducks doing?

- ⓘ Flying.
- ⓙ Cheeping.
- ⓚ Calling.
- ⓛ Sleeping.

34. What is a good title for this passage?

- Ⓜ Early Morning.
- Ⓝ The Eastern Sky.
- Ⓞ A Dark Day.
- Ⓟ Winter.

GO ON

A fly has two big cup-shaped eyes, one on each side of its head. Each eye is made up of hundreds of tiny eyes. Each of the tiny eyes can see. However, each little eye does not see everything that is around the fly. Each little eye sees only what is right in front of it. Because the eye is cup-shaped, each little eye faces in a different direction. So each little eye sees a different part of what is around the fly.

35. Each tiny eye sees what is

- Ⓐ all around it.
- Ⓑ in front of it.
- Ⓒ to the left of it.
- Ⓓ to the right of it.

36. How many tiny eyes does each big eye have?

- Ⓐ A hundred.
- Ⓑ More than a hundred.
- Ⓒ One.
- Ⓓ Two.

37. Because of the big eye's shape, each tiny eye

- Ⓐ can look at something different.
- Ⓑ can move.
- Ⓒ can close.
- Ⓓ is cup-shaped.

If you walk much in the woods, you may someday meet Mrs. Grouse with her brood. She will probably rush at you and hiss, or she may drag one wing on the ground and pretend she is hurt. Meanwhile, the babies have hidden under leaves and branches. With this trick the mother hopes to lure you away from them. When all is safe again, she signals the young and they scramble back to her side.

38. The mother doesn't try to hide because she

- Ⓐ has been hurt.
- Ⓑ knows there is no danger.
- Ⓒ thinks you can't see her.
- Ⓓ wants you to notice her.

39. You can tell from the passage that these babies usually

- Ⓐ leave their home only at night.
- Ⓑ stay close to their mother.
- Ⓒ cannot see or hear well.
- Ⓓ rush at you and hiss.

40. All is safe again when

- Ⓐ the babies have hidden.
- Ⓑ the babies have left.
- Ⓒ you have left.
- Ⓓ night has come.

GO ON

One spring day, a five-year-old Navajo boy named Fred Bia was watching the family sheep flock. It was his daily chore to follow the sheep as they drifted over the red, rocky earth in their endless search for grass and the leaves of semi-desert plants. He had covered this ground so many times that he no longer paid any attention to where he was. His thoughts wandered as he moved slowly with the animals.

When he saw the rock in front of him, he knew he was in a place that he had not been before. He could not believe his eyes. The big red rock was covered with drawings of people and animals. Fred stood very still as he stared at them. An excitement he had never felt before raced through his blood. Who had made these pictures? When? He had no idea.

41. The day was unusual because of what Fred

- (A) heard.
- (B) found.
- (C) believed.
- (D) had planned to do.

42. What was Fred doing while he walked?

- (E) Singing.
- (F) Looking all around.
- (G) Looking for semi-desert plants.
- (H) Thinking about things.

43. Fred could tell he was in a new place because of the

- (I) rock.
- (J) sheep.
- (K) grass.
- (L) stream.

44. Why did Fred stand very still?

- (M) He didn't want to frighten the sheep.
- (N) He was lost.
- (O) He was too tired to run away.
- (P) He was surprised.

45. Fred asked himself how

- (Q) the drawings got there.
- (R) to get home.
- (S) to find more grass.
- (T) many of the sheep were left.

TASK 2: Additional Below-Grade-Level Multiple-Choice Questions

Space ships in movies are really just little models of space ships. One movie had a very special camera. It was used to film the little models. It made them look much bigger than they were. And it also made them look as if they were flying.

As the camera took pictures it kept moving. Movie people say it was "traveling." Sometimes it went away from a model, backward. Sometimes it went diving under one. On the screen it looked as if the model were flying and climbing.

49. This story is about space ships in

o museums.
o space.
o battle.
o films.

50. These space ships are really

o pictures.
o models.
o people.
o airplanes

51. The space ship seemed to climb because

o it looked much bigger.
o it was "traveling."
o it was upside-down.
o the camera dived.

A little bear lived on an iceberg. One day, he saw a big rainbow in the sky. He thought what fun it would be if he could change his white coat into a colored one.

52. The bear wanted to change the color of

o the sky.
o the snow
o his coat
o his house

53. What gave the bear his idea?

o A rainbow.
o A cloud.
o A friend.
o An iceberg.

The Window

Behind the blind I sit and watch
The people passing - passing by;
And not a single one can see
My tiny watching eye.

They cannot see my little room,
All yellowed with the shaded sun,
they do not even know I'm here;
Nor'll guess when I am gone.

54. What makes the room look yellow?

o The child has one eye closed.
o The paint is old.
o A lamp is on.
o The sun is shining through a shade.

55. The child thinks he is

o gone.
o with a friend.
o sick.
o hidden.

Some parts of Africa have two kinds of summer - the rainy summer and the dry.

During the rainy summer, the forests are green with plants. On the plains, the grass grows tall. There are streams and water holes everywhere.

Then the rains stop. Day after day a bright sun shines. Most of the streams dry up.

The only water for miles around is the big river. To it, all the animals must come to drink.

56. Why do the animals have to go to the big river?

o They need to cross it.
o They need water to drink.
o It is green with plants.
o People drive them there.

57. When does the wet summer come?

o Right before the fall.
o Right after the fall.
o Right before the dry summer.
o Right after the dry summer.

58. This story is mainly about African

o weather
o rivers
o animals
o plants

Momo is the name of a little girl who was born in New York. The word Momo means "the peach" in Japan, where her father and mother used to live. On her third birthday Momo was given two presents - red rubber boots and an umbrella. They pleased her so much that she even woke up that midnight to take another look at them.

59. What does Momo mean in Japanese?

o Mama.
o Little girl.
o The peach.
o Birthday

60. Why did Momo look at her gifts again?

o she forgot what they were.
o She like them.
o She was in New York.
o She didn't want to go to bed.

61. The story says that Momo took another look at her gifts

o under the bed.
o during the night.
o at the party.
o in the morning.

The Siamese Fighting Fish has an unusual nest. The father makes it. It is a bubble nest. It floats at the top of the water. The mother lays her eggs. Father catches them in his mouth. He blows them into the nest. Then he chases mother away. She leaves. But father stays. He watches the nest. He keeps it safe. Finally the babies hatch. But they can not swim yet. They must stay in the nest for three more days. Sometimes a baby falls out. Then the father hurries to it. He takes it in his mouth. He blows it back into the nest.

62. This fish's nest is made by the

o babies.
o leaves.
o father
o mother.

63. After the mother lays her eggs, she

o goes away.
o watches them.
o cleans them.
o catches them.

Samuel Duck wanted to build a tree house. "I know just the place for one," he said. "I can build it all by myself, too." Samuel hurried to the woodshed for a ladder. He carried it over to the big apple tree, and up he went.

"Oh, do be careful!" called Samantha from below. "Ducks don't belong in trees."

64. Why did Samuel get a ladder?

o To get to the roof.
o To climb a tree.
o To go below.
o To reach Samantha.

65. Where was Samantha?

o Under the tree.
o Up in the tree.
o In the woodshed.
o Across the street.

66. Samantha was worried because she thought Samuel

o might get hurt.
o belonged in trees.
o might get in trouble.
o would break the tree.

Sneakers was a rapscallion cat from the time he was born. The other kittens stayed in the box until they grew up enough. But Sneakers was off even before he grew up enough.

Sneakers went after little black bugs. He chased butterfly shadows across the ground. He pounced on people's shoelaces if they were the least bit hanging.

67. The first sign that Sneakers was a rapscallion was his

o getting out of the box.
o being born.
o pouncing on shoelaces.
o fighting.

68. What did Sneakers like to do?

o Hide.
o Stay with the other kittens.
o Try to catch things.
o Make friends with black bugs.

Mexican American settlers made houses out of what the land had to offer - adobe, or heavy clay.

Adobe was molded by mixing the clay with straw and pouring it into molds. The adobe dried in the sun, and it took the shape of a brick. Adobe bricks permitted people to build high thick walls. The Mexicans added windows, fireplaces, and stairs to the homes. The thick adobe brick walls kept our the heat in summer and the cold in winter.

69. What was added to the adobe?

o Clay.
o Paint.
o Straw.
o Bricks.

TASK 3: Short-Answer Literal and Inferential Questions (Below-Grade-Level Items: #'s 1-27; Grade-Appropriate Items: #'s 28-59)

Hal was sitting on the front steps. His brother came out and asked him if he wanted to go to the movies.

1. Where did Hal's brother find him?

_____

_____

_____

Gwen thought she heard a truck. But it was really a plane overhead

2. What made the noise she heard?

_____

_____

_____

Kris couldn't hear over the phone. The TV was too loud.

3. So what did Kris do?

_____

_____

_____

Yoshie was late for school today. She was still hurrying down the hall when the last bell rang. The class had already started when she got to her seat.

4. Where was she when the bell rang?

_____

_____

_____

Mary brushed her hair. She put on her socks and pants and shirt. Then she looked in the closet. "Mother!" she called. "Where are my shoes?"

5. What had Mary lost?

_____

_____

_____

It was spring in Peapack. The mayor called a meeting of the town. "Contest at the county fair!" he shouted. "And a cash prize to the town that grows the finest vegetables!"

6. Why did the mayor shout?

_____

_____

_____

If you look carefully at the front legs of an ant, you will see one place with long stiff hairs. These hairs are used like a brush. An ant lifts its leg up and pulls an antenna through the "brush." An ant also licks clean every part of its body it can reach with its tongue. Any part the tongue can't reach, the ant cleans with its feet.

7. An ant has a brush on

_____

_____

_____

8. An ant uses the "brush" to clean

_____

_____

_____

A turtle lay on a rock. Because of the color of her shell she looked very much like the rock. She lay in the sun and waited for a mayfly. A mayfly flew down. The turtle stretched her long neck out. With a quick snap of her jaws - the mayfly was gone. The turtle laid her head on one foot and waited for another fly to come close.

9. Why did the turtle look like the rock?

_____

_____

_____

10. While she was waiting, the turtle was being

_____

_____

_____

11. What did the turtle want to do?

_____

_____

_____

Silver is a metal. It is shiny. And it is soft for a metal.

Mix one part copper with nine parts silver. You have sterling silver. It is harder than pure silver. It makes beautiful knives, forks and spoons. Silver makes jewelry, too. But it gets dark easily. We say it tarnishes. Then it must be polished.

12. Compared to most other metals, silver is

_____

_____

_____

13. How does silver look when it tarnishes?

_____

_____

_____

Saturday morning is jump-off time. Aunt Martha and I are going to drive all the way to North Carolina in her new car.

Aunt Martha says, "No boys and no men, just women."

We made our list last week and double-checked it every night. Aunt Martha forgets things, so I'm her reminder.

Last year she forgot the maps and our lunch on the kitchen table.

14. Jump-off time is when Aunt Martha will

_____

_____

_____

15. Thinking about the trip makes the person who is telling the story feel

_____

_____

_____

Evan went to the playground. He took his toothbrush glass and a spoon.
The paving of the playground was cracked. Grass and weeds grew up through the broken concrete.
Evan found a weed that had big, lacy flowers on it. He dug it up with his spoon. He planted it in his toothbrush glass.

Then he took it home.

16. What was the playground covered with?

_____

_____

17. What did Evan go to the playground to look for?

_____

_____

_____

18. What was the playground like?

_____

_____

_____

19. At the playground, Evan used his spoon for

_____

_____

_____

Sneakers was a rapscallion cat from the time he was born. The other kittens stayed in the box until they grew up enough. But Sneakers was off even before he grew up enough.

Sneakers went after little black bugs. He chased butterfly shadows across the ground. He pounced on people's shoelaces if they were the least bit hanging.

20. When did Sneakers start being a rapscallion?

_____

_____

_____

Ramon made a pencil holder for his father. He began by painting an empty can. Then he cut out pictures of animals and pasted them on the can.

21. What did Ramon do first?

_____

_____

_____

One bright, sunny morning, Allison was playing in front of her house when she saw a truck loaded with furniture. "Please stop next door," Allison said. she crossed her fingers. There was a loud crashing "bang," and the truck stopped. Allison crawled under the hedge. her favorite hideout. Hours seemed to creep by while she played with her paints and watched the moving men unload the truck. At last Allison saw a bicycle. There must be *someone* for me to play with. she thought.

22. Why did Allison cross her fingers?

_____

_____

_____

23. Where did Allison like to hide?

_____

_____

_____

24. What was Allison hoping for?

_____

_____

_____

25. Why did Allison go under the hedge?

_____

_____

_____

Rafer woke up with that good feeling. It was his birthday. a merry mellow me day. There would be cake with supper. He could pick the TV shows. And no chores.

Birds were singing on the fire escape. "Birds. I love you," he whispered. They flew away.

26. Where were the birds when Rafer heard them?

_____

_____

_____

27. You can tell from the story that Rafer often had to

_____

_____

_____

Whales are not fish. They belong to the same group of animals as cows, lions, dogs, and people. Whales are mammals and like all mammals, they breathe air. They are warm-blooded, bear their young alive, and nurse them with milk. All mammals are hairy. But the dark gray or blackish skins of whales have no hair. Only a few whiskers remain to show that the ancestors of whales were once hairy animals. These distant ancestors of whales were land animals.

28. In what way are whales like lions?

_____

_____

_____

29. Where did the whales' ancestors live?

_____

_____

_____

The first railroad across Canada took many years to build. For months and years, engineers and workmen struggled to lay tracks through the forests of northern Ontario and around the lakes and over the swamps that lie between the Great Lakes and the Red River. The tracks kept sinking in the swamps. In one place several lots of track sank before the engineers found firm ground. The work was easier on thePrairies. But then there were the Rockies! Here, tunnels had to be blasted through the mountains and bridges build across the rivers.

30. Where did the work go faster?

_____

_____

_____

31. Why did the tracks sink?

_____

_____

_____

32. After the workers had crossed the prairies, what make their work hard?

_____

_____

_____

Sky Seasoning

A piece of sky
Broke off and fell
Through the crack in the ceiling
Right into my soup.
KERPLOP!
I really must state
That I usually hate
Lentil soup, but I ate
Every drop!
Delicious delicious
(A bit like plaster).
But so delicious, goodness sake-
I could have eaten a lentil-soup
lake.
It's amazing the difference
A bit of sky can make.

33. What went kerplop?

_____

_____

_____

34. The poem suggests that there is something wonderful about

_____

_____

_____

35. Why could the poet have eaten a whole lake?

_____

_____

_____

A forest ranger was sleeping in his tent one night when something touched his lips softly. Smiling in his sleep, he brushed it aside. The caress was repeated. It felt cold and a bit <u>prickly</u>. The ranger opened his eyes in pitch darkness. Reaching for his flashlight, he flooded the tent with light. There, looking wistfully down on him, were two mountain lions!
No one will ever know who got our of the tent first, the lions or the ranger- but the ranger kept on going.

36. What made the ranger wake up?

_____

_____

_____

37. What felt <u>prickly</u>?

_____

_____

_____

38. When the ranger saw the lions, he

_____

_____

_____

Jennifer Allen chuckled to herself as she darted around a corner and raced down the narrow corridor. She could outwit Jim and Susan any time even though she was their younger sister. They were after her to go swimming with them in the pool down below. but Jennifer wanted to be alone. She had spent hours exploring all the nooks and crannies of this big ship that was taking them to Canada. And she had come upon a secret place: a place where she could spin her dreams without being disturbed.

39. Where was Jennifer?

_____

_____

_____

40. Why did Jennifer chuckle?

_____

_____

_____

The men were great hunters in the village of Temlaham. Every spring and fall they hunted goats that roamed the highest peaks of the nearby mountain, Stek-yaw-den. They killed only what was needed for meat and skin and horn. according to the old law.

41. What was the old law about?

_____

_____

_____

42. Where did the men hunt the goats?

_____

_____

_____

Little

I am the sister of him
And he is my brother
He is too little for us
To talk to each other

So every morning I show him
My doll and my book
But every morning he is still is
Too little to look.

43. What is he?

_____

_____

_____

44. What does the poet wish he would do?

_____

_____

_____

A Sioux leader was talking.
"In the buffalo days," he said, "we were happy here in our own country. The prairie was full of four-legged game. The young men hunted. We were never hungry, unless the winter was long and the snow deep. Then the Wasichus came. These settlers told us they wanted only a little land. Enough to set a house on. But they kept coming, like a river. They filled the prairie. Their guns drove the buffalo far away. Now our young men have little to hunt. Nothing to do."

45. Where was our own country?

_____

_____

_____

46. What made the settlers seem like a river?

_____

_____

_____

47. Why don't the young men have anything to do?

_____

_____

_____

The harvest mouse is a marvelous acrobat. His light weight means he can jump from one slender corn stalk to another without breaking them. He uses his long tail to help him balance, and also as a brake when climbing down stalks. The outer toe on each back foot is used for gripping, like a thumb. The front feet, like hands, can gather and hold food.

48. The harvest mouse doesn't break the corn stalks because it is

_____

_____

_____

49. What does the harvest mouse use its tail for?

_____

_____

_____

50. One toe on each back foot is like

_____

_____

_____

51. This story is mainly about why the mouse is so good at

_____

_____

_____

Most whales come up for air about every seven or eight minutes. Some whales may stay down for as long as an hour.
Whales breathe through blowholes in the top of their heads. Blowholes are tightly shut when whales are underwater. They open only when whales come up to breathe.
Whales blow warm air out through the blowholes. The warm air looks like fog when it hits the cold air.
A whale does not breathe through its mouth. There is no opening between the throat and the lungs of a whale. The mouth is for eating.

52. A whale gets air through

_____

_____

_____

53. When the whale goes under the water, its blowhole

_____

_____

_____

One morning, Emma's father announced that their cow, Blackie was going to have a calf. Emma was so excited. Each day she checked on Blackie and when it seemed as though she'd been waiting forever, her father called the whole family out to the barn. There, lying curled up in a pile of fresh straw, was a fine, healthy calf.

Blackie stood there licking her calf dry and making gentle sounds. Then, everyone watched the calf as he struggled to stand. He was very wobbly, but he managed to get up, first on his hind legs, then on his front legs, until he stood trembling beside his mother. Emma turned to her brother and sister and said proudly, "Let's call him Little Blackie?" and that is just what they did.

54. What was hard for the calf to do?

_____

_____

_____

55. What was Blackie doing to the calf?

_____

_____

_____

56. Who named the calf?

_____

_____

_____

57. How did the calf get up?

_____

_____

_____

A pueblo child tells how a desert plant is used.

Our Mothers gather
The flowers of guaco,
And the leaves and stems;
They boil them together
On a fire in the plaza;
For many days
They boil the guaco.

They make little, soft cakes
Of the boiled-down guaco,
And wrap them
In corn-husk blankets
To dry into paint,
Into black paint,
To make the designs
On pottery.

58. What is the paint made from?

_____

_____

_____

59. What is the paint used on?

_____

_____

_____

Appendix B: Explanation of how readability levels were determined for the items on the Gates-MacGinitie Reading Comprehension subtest:

Readability levels were determined by the test authors using three readability formulas: Dale-Chall (Dale & Chall.1968), Fry (Fry, Fountoukidis, & Polk, 1985), and Harris-Jacobson (Harris & Jacobson, 1982). "Readability estimates from at least two of the three formulas were averaged to obtain an average readability figure for each prose passage. The Fry formula is intended for rating materials for Grade 1 though college level, so Fry readability estimates were included in all averages. Since the original Dale-Chall formula had a floor of about Grade 4, and since the new formula is based on an updated version of the same word list, Dale-Chall readability estimate below 4.0 were not included in the averages (MacGinitie & MacGinitie, 1989, p. 8-9)."

Appendix C: Definitions of Literal and Inferential Questions and Examples for Graduate Student Judges

Literal and Inferential items:

Please rate the questions following each passage as **L** for literal or **I** for inferential based on the following definitions and examples:

1. **Literal**: An item should be considered ***literal*** if:

the student can answer by choosing something stated explicitly in the text or choosing a restatement of something stated explicitly in the passage.

Example:
Passage:

It was a good thing for Neesie that Mama couldn't see her, or she would have got a good spanking. Mama couldn't hear her either, but I could. All the time Mama was cornrowing my hair, Neesie kept calling me and waving her arms around, trying to make me look at her. After a while, I got tired of it. "Stop it Neesie!" I said.

Literal question:
What was Mama doing?
a) Waving her arms.
b) Yelling at Neesie.
c) Fixing the girl's hair (correct answer: it's a restatement of "cornrowing my hair")
d) Getting dinner ready

2. **Inferential:** An item should be considered ***inferential*** if:

the student can *not* answer the question by choosing a restatement of something stated explicitly in the passage. The reader needs to apply knowledge of the world to produce information or make connections that are not explicitly stated in the text.

Example:
Same passage as above:

Inferential Question:
How was Neesie acting?
a) tired
b) polite
c) naughty (correct answer: inferential, because the reader must make connection that Neesie was doing something that was deserving of a spanking)
d) surprised

Appendix D: Answers to Short-Answer Questions:

Question:	**List of acceptable responses**
1. Where did Hal's brother find him?	On the front steps, on the steps/stairs, sitting on the front steps
2. What made the noise she heard?	A plane, a plane overhead
3. So what did Kris do?	Turned the TV/volume down, turned off the TV
4. Where was she when the bell rang?	In the hall, hurrying down the hall, running down the hall
5. What had Mary lost?	Her shoes
6. Why did the mayor shout?	So everyone could hear him, he was excited
7. An ant has a brush on	Its front legs, its legs
8. An ant uses the brush to clean	Its antenna
9. Why did the turtle look like the rock?	Because of the color of her shell, because her shell looked like the rock, she was the color of the rock
10. While she was waiting, the turtle was being	Still, quiet, sneaky, silent, camouflaged, patient
11. What did the turtle want to do?	Catch a mayfly/fly, eat a mayfly/fly
12. Compared to most other metals, silver is	Soft
13. How does silver look when it tarnishes?	It gets dark, dark, it's a dark color, black, brown
14. Jump-off time is when Aunt Martha will	Drive to N. Carolina, go to N. C., go on a road trip/trip
15. Thinking about the trip makes the person who is telling the story feel	Happy, good, excited, glad
16. What was the playground covered with?	Grass and weeds, concrete, paving, pavement, cement
17. What did Evan go to the playground to look for?	Weeds, a big weed, a plant, a flower (to bring home), a flower
18. What was the playground like?	Old, it had cracked concrete/pavement, covered with grass and weeds, weedy, broken concrete
19. At the playground, Evan used his spoon for	Digging up a weed, digging, a shovel
20. When did Sneakers start being	

rapscallion?	From the time he was born, when he was very little, when he was a baby
21. What did Ramon do first?	Painted an empty can, painted a can, painted
22. Why did Allison cross her fingers?	She wanted the truck to stop next door, she wanted a new next-door neighbor
23. Where did Allison like to hide?	Under the hedge/bush
24. What was Allison hoping for?	A friend, someone to play with, that the truck would stop
25. Why did Allison go under the hedge?	To watch the moving men unload the truck, to spy on the men, she didn't want to be seen
26. Where were the birds when Rafer heard them?	On the fire escape
27. You can tell from the story that Rafer often had to	Do chores/jobs
28. In what way are whales like lions?	They are mammals, they breathe air, they belong to the same group, they are warm-blooded, they bear young live, they nurse their young with milk
29. Where did the whales' ancestors live?	On land
30. Where did the work go faster?	On the prairies
31. Why did the tracks sink?	Because they were in the swamps, because the ground wasn't firm, because the ground was soft
32. After the workers had crossed the prairies, what made their work hard?	The Rockies, the mountains, the rivers, tunnels had to be blasted, bridges had to be built
33. What went kerplop?	A piece of sky, the sky falling into the soup
34. The poem suggests that there is something wonderful about	A piece of sky, the sky, lentil soup with a piece of sky
35. Why could the poet have eaten a whole lake?	The soup was so delicious/good/yummy/tasty, it was so good, he liked the soup so much
36. What made the ranger wake up?	Something touched his lips, the lions touched him, the lion's whiskers, a prickly feeling
37. What felt prickly?	The lions, the lions' whiskers/tongue, the touch/caress
38. When the ranger saw the lions, he	Ran out of the tent, ran away, was surprised/scared
39. Where was Jennifer?	In a narrow corridor, on a ship
40. Why did Jennifer chuckle?	Because she could outwit/outsmart Jim and Susan/her brother and sister, because she got away from Jim and Susan/her brother and sister

41. What was the old law about?	Kill only what is needed for meat and skin and horn, only kill what you need
42. Where did the men hunt the goats?	On the highest peaks of the nearby mountains, on the mountain, on Stek-yaw-den
43. What is he?	Her brother, her little brother, a baby
44. What does the poet wish he would do?	Play with her, look at her doll/her book, talk to her, grow up/be older/get bigger
45. Where was our own country?	On the prairie
46. What made the settlers seem like a river?	They kept coming
47. Why don't the young men have anything to do?	There are no buffalo to hunt, the buffalo were gone, they have little to hunt
48. The harvest mouse doesn't break the corn stalks because it is	Light, lightweight
49. What does the harvest mouse use its tail for?	To help him balance, a brake
50. One toe on each back foot is like	A thumb, a gripper
51. This story is mainly about why the mouse is so good at	Jumping, climbing, acrobatics
52. A whale gets air through	Its blowhole
53. When the whale goes under the water, its blowhole	Shuts tightly, closes, shuts
54. What was it hard for the calf to do?	Stand up/stand
55. What was Blackie doing to the calf?	Licking it dry, licking it
56. Who named the calf?	Emma
57. How did the calf get up?	On his hind legs and then on his front legs, on wobbly legs slowly, carefully, it struggled
58. What was the paint made from?	The flowers/leaves/and stems of guaco, guaco
59. What is the paint used on?	Pottery

VITA

Laura Brueggeman Green
342 N. 82nd St.
Seattle, WA 98103
(206) 297-1127

EDUCATION	Doctoral Candidate in Speech and Hearing Sciences, October, 1997, University of Washington, Seattle, Washington
	Master of Arts in Communication Disorders, May 1990, University of Houston, Houston, Texas
	Bachelor of Science in Speech/Language Pathology December, 1987, Stephen F. Austin State University, Nacogdoches, Texas
PROFESSIONAL EXPERIENCE	
September 1999-present	Research Coordinator: Multidisciplinary Learning Disabilities Center, Department of Educational Psychology, University of Washington, Seattle, WA. Responsibilities include managing research assistants, conducting statistical analyses on project data, writing research articles, planning collaborative meetings with teachers from local school districts, and completing general administrative tasks.
	Continued management of the UW Language and Literacy Clinic.
July, 1997-August, 1999	Pre-Doctoral Research Assistantship II: University of Washington, Seattle, WA. Responsibilities included gathering and scoring classroom and individual student assessment data, managing all grant databases, and conducting statistical analyses for the Multidisciplinary Learning Disabilities Center in the Department of Educational Psychology.
	Continued management of the UW Language and Literacy Clinic.
August, 1995-May, 1997	Pre-Doctoral Teaching Assistantship I: University of Washington, Seattle, WA. Various quarterly responsibilities included teaching SPHSC 100, a speech class for non-majors; serving as a teaching assistant for SPHSC 405, a diagnostic methods class; and serving as a clinical supervisor in the University of Washington Speech and Hearing Clinic.
	Created, managed, and conducted clinical supervision in the Language and Literacy Clinic, an assessment service in the University of Washington Speech and Hearing Clinic.
August, 1992-August, 1995	Visiting Faculty: University of Houston, Houston, TX.

Responsibilities included teaching undergraduate courses in normal language development and pre-school language disorders, and graduate courses in diagnostic methods in speech/language pathology and school-age language disorders. Responsibilities also included supervision of a graduate student diagnostic team in the School of Optometry's Interdisciplinary Clinic (assessment of language/learning disabilities), and summer clinical supervision of diagnosis and treatment.

Private Practice: Houston, TX. Responsibilities included assessment and treatment of school-age children with language/learning disabilities.

June-August, 1992 — Clinical Supervisor: University of Houston, Houston, TX Responsibilities included observation of clinical treatment sessions and providing student clinicians with verbal and written feedback.

August, 1990 - June, 1992 — Speech/Language Pathologist: Spring Branch Independent School District, Houston, TX Responsibilities included the diagnosis and treatment of elementary and high school students in the areas of articulation, augmentative communication, phonological processing, language/learning disabilities, language delays, fluency disorders, and Community Based Instruction

CERTIFICATION — American Speech and Hearing Association/Certificate of Clinical Competence in Speech/Language Pathology

PROFESSIONAL AFFILIATIONS

1990 - present — American Speech and Hearing Association

1992-1995 — Texas Speech and Hearing Association
Houston Association of Communication Disorders (Secretary)

PUBLICATIONS

McCutchen, D., Abbott, R., Green, L., Beretvas, S., Potter, N., Cox, S., Gray, A. Beginning literacy: Links among teacher knowledge, teacher practice, and student learning. *Journal of Learning Disabilities*, in press.

Klecan-Aker, J. & Green, L. Language treatment of a child with sensory integration deficit. *Journal of Child Language Teaching and Therapy*, 1994.

Klecan-Aker, J. and Brueggeman, L. *The Expression Connection: A Structured Approach to Teaching Storytelling to School Age Children.* The Speech Bin, Vero Beach, Fla. 1991.

STATE AND NATIONAL PRESENTATIONS

Green, L. and McGehee, K. *Literal and Inferential Comprehension in Fourth Grade Readers: Effects of Readability and Answer Format.*

Poster Presentation. American Speech and Hearing Association Convention, San Francisco 1999.

Green, L. and McGehee, K. *Teaching Reading Using a Systematic, Functional Approach: A Case Study*. Poster Presentation. American Speech and Hearing Association Convention, San Antonio 1998.

Klecan-Aker, J. & Green, L. *Sensory Integration and Language Learning Disabilities: A Treatment Case Study*. Poster Presentation. American Speech and Hearing Association Convention, Orlando 1995.

Klecan-Aker, J. & Brueggeman, L *Language Treatment of School Age Children: A Look at Efficacy and Outcomes* Miniseminar. Texas Speech and Hearing Association Convention, Houston 1995.

Klecan-Aker, J. & Brueggeman, L. *Narratives in School Age Children: Theoretical Issues, Assessment and Treatment*: Miniseminar. American Speech and Hearing Association Convention, Anaheim 1993.

Brueggeman, L. & Klecan-Aker, J. *Structure in Teaching Storytelling: A Treatment Program and Its Effects* Poster Session. American Speech and Hearing Association Convention, San Antonio 1992

Klecan-Aker, J. & Brueggeman, L. *Assessment and Treatment of a School Age LLD Child: A Case Study* Poster Session. American Speech and Hearing Association Convention, Seattle 1990.

Brueggeman, L. *Assessment and Treatment of a School Age LLD Child: A Case Study* Miniseminar. Texas Speech and Hearing Association Convention, Dallas 1990

SERVICE (PRESENTATIONS)

Standard and Criterion-Referenced Testing: Diagnosis of Language and Learning Disabilities in School-Age Children, Houston Association of Communication Disorders Spring Workshop, 1995.

Phonological Processing Assessment and Treatment in the Schools, Spring Branch Independent School District Workshop, Houston, Spring, 1992.

Treatment of Narrative Difficulties in Elementary School Children Spring Branch Independent School District Workshop, Houston, Fall, 1991.