

ABSTRACT

Title of Document: EFFECTS OF TEXT MARKERS AND
FAMILIARITY ON COMPONENT
STRUCTURES OF TEXT-BASED
REPRESENTATIONS

Marcia H. Davis, Doctor of Philosophy, 2006

Directed By: Professor John T. Guthrie
Department of Human Development

Reading theorists agree that the outcome of reading comprehension is a text representation (Gernsbacher, Varner, & Faust, 1990; Kintsch, 1998). To measure reading comprehension, however, many use testing formats such as multiple-choice and short answer, that have been shown to provide very little information about the text representations created during reading (Kintsch & Kintsch, 2005). A different type of format, proximity measures, is a promising measure for text based representations, although few studies have examined the validity of this type of comprehension measure. The current dissertation addressed this issue by creating a proximity measure named the Passage Comprehension for Structured Text (PCST) and by examining the validity of the PCST through experimental manipulations on the text.

This investigation tested the comprehension of 236 ninth-grade students with the PCST. Students were asked to read a short text followed by a computer task

where the students rated the similarity of 11 concepts. Two components were extracted from the PCST, the textbase component and the situation model component. Text manipulations included changes in text coherence and familiarity. It was hypothesized that a coherence manipulation should have an effect on the textbase component and a familiarity manipulation should have an effect on the situation model component. Further, both manipulations should influence the strength of the factor that determines the scores on these components.

A multivariate analysis of variance was used to compare the conditions. Results confirmed that students with coherent text outperformed students with incoherent text on the textbase component and students with familiar text outperformed students with unfamiliar text on the situation model component. Confirmatory factor analysis was used to further explore the effect of text manipulations on the structure of the PCST components. Results indicated that there was a stronger factor for the situation model component when the text was familiar compared to when the text was unfamiliar. Limited evidence suggests that there was also a stronger factor for the textbase component when the text included macrosignals compared to when the text did not include macrosignals.

EFFECTS OF TEXT MARKERS AND FAMILIARITY ON COMPONENT
STRUCTURES OF TEXT-BASED REPRESENTATIONS

By

Marcia H. Davis

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2006

Advisory Committee:
Professor John T. Guthrie, Chair
Professor Peter P. Afflerbach
Professor C. Mitchell Dayton
Professor Mariam J. Dreher
Professor Allan Wigfield

© Copyright by
Marcia H. Davis
2006

Acknowledgements

I dedicate this dissertation to my wonderful husband, Randy, and to our beautiful daughter, Julie. You are my greatest supporters. Thank you for your optimism and encouragement and for helping me to keep my priorities in check. I love you very much.

I would like to thank Dr. John Guthrie for guiding me through the dissertation process. You have a special talent for identifying and helping students build on their strengths. Without your expertise, mentoring, and patience, I would not have achieved my goals. I would like to also thank the other members of my dissertation committee, Drs. Allan Wigfield, Chan Dayton, Jean Dreher, Roger Azevedo, and Peter Afflerbach, for their additional support. Your guidance and thoughtful questions allowed me rethink and extend my dissertation study in new ways.

I would like to thank Eileen Kramer and Ellen Kaplan for providing friendly feedback on grammar and APA style. I would additionally like to thank Eileen Kramer for believing in me and encouraging me especially when my confidence was shaken. You have helped me in numerous ways and have become a great friend.

Special thanks are in order for my parents, Anita and Daniel Hardisky. Thank you for your love and support. You have always encouraged me to follow my dreams. I am also very grateful to my mother-in law, Mary Riniker, for her help in obtaining subjects and collecting data for this dissertation.

Finally, I would like to thank my fellow graduate students, Drs. Ana Taboada, Stephen Tonks, and Kathleen Perencevich for sharing their experiences of the

dissertation process. Your guidance, openness, and support made the dissertation process easier than it would have otherwise been.

Table of Contents

Acknowledgements.....	ii
Table of Contents.....	iv
List of Tables.....	vi
List of Figures.....	vii
Chapter 1: Introduction.....	1
Psychometric Studies on the Multidimensionality of Reading Comprehension.....	2
Reading Theory.....	4
Theories of Reading Comprehension.....	4
Coherence Effects on Text Comprehension.....	6
Background Knowledge Effects on Text Comprehension.....	8
Measurement of Text Representations.....	9
Pilot Investigation.....	10
Purposes and Significance of the Dissertation Investigation.....	11
Chapter 2: Review of the Literature.....	20
Purpose of Literature Review.....	20
Psychometric Studies on the Multidimensionality of Reading Comprehension.....	22
Independent Skills Theory.....	24
Global Skills Theory.....	33
Hierarchical Skills Theory.....	37
Theories of Reading Comprehension.....	41
Construction-Integration Theory.....	43
Grimes’s Text Representation Formation.....	62
The Structure Building Framework.....	63
Measuring Organized Knowledge Built from Text.....	65
Variability of Text: Coherence and Background Knowledge.....	73
Research on Increasing Coherence of Text.....	73
Increasing Coherence Through use of Macrosignals.....	76
Increasing Coherence Through use of Text Structure Strategies.....	81
Research on the Influence of Background Knowledge.....	84
Theoretical Expectations and Hypotheses.....	90
Expectations.....	90
Conclusion.....	96
Theoretical and Practical Significance.....	96
Hypotheses.....	97
Chapter 3: Method.....	99
Research Hypotheses.....	99
Participants.....	99
Materials.....	100
Measure: Passage Comprehension for Structured Text.....	105
Definition.....	106
Procedure.....	106
Validity and Reliability.....	109

Scoring	111
Administrative Procedures.....	113
Chapter 4: Results.....	115
Preliminary Analyses.....	115
Descriptive Statistics.....	115
Reliability Analysis.....	119
Multivariate Analysis of Variance.....	119
Research Hypothesis 1.....	120
Research Hypothesis 2.....	120
General Structure Analysis.....	121
Control Group Analysis.....	121
Confirmatory Factor Analyses.....	122
Research Hypothesis 3.....	122
Research Hypothesis 4.....	125
Further Analysis: The Two Factor Model.....	126
Factor Loadings for Separate Models.....	133
Research Hypotheses 3 and 4 Assuming Ordinal Data.....	133
Chapter 5: Discussion.....	138
PCST as a Measure of Knowledge Representation.....	139
PCST Sensitivity to the Macrosignals Manipulation.....	140
PCST Sensitivity to Familiarity Manipulation.....	144
Control Analyses.....	148
The Salience of the Textbase and Situation Models.....	150
Effect of Macrosignals on the Salience of the Factors.....	151
Effect of Familiarity on the Salience of the Factors.....	153
Additional Analyses on the Fit of Textbase and Situation Model.....	157
Limitations.....	160
Future Directions.....	163
Appendix A.....	165
Appendix B.....	166
Appendix C.....	167
Appendix D.....	168
References.....	187

List of Tables

1. Definitions of terms.....	17-19
2. Expert ratings for each text	111
3. The number of students in each condition before equating the groups	116
4. Means and standard deviations of textbase and situation model connections for each text	118
5. Intercorrelations among familiarity, macrosignals, situation model connections, and textbase connections in the treatment groups.....	118
6. Model fit for models 1, 2, and 3 in the textbase and situation model analyses.....	131
7. Model fit for separate models	132
8. Factor loadings for separate models.....	133
9. Model fit for models 1, 2, and 3 in the textbase and situation model analyses with ordinal data.....	136
10. Model fit for separate models with ordinal data.....	137
11. Means and standard deviations of textbase and situation model connections for each text.....	175
12. Intercorrelations among familiarity, macrosignals, situation model connections, and textbase connections.....	176

List of Figures

1. Hierarchical structure of <i>The Birth of a Baby Shark</i>	54
2. Representation of the text structure for America.....	104
3. Representation of the text structure for Sri Lanka.....	104
4. Representation of the two-factor model.....	127
5. Representation of the text structure for America (Pilot).....	172
6. Representation of the text structure for Sri Lanka (Pilot).....	172
7. Passage comprehension computer screen	174
8. Student textbase scores given reading macrosignals	178
9. Student situation model scores given reading macrosignals	179
10. Student situation model scores given topic familiarity	180
11. Student textbase scores given topic familiarity	181

Chapter 1: Introduction

The complexity of the reading comprehension process is well known (Ferstl & Kintsch, 1999; Gernsbacher, Varner, & Faust, 1990; Kintsch & Kintsch, 2005). This complexity is due to the diverse processes occurring during reading as well as the characteristics of the text and the reader. Likewise, a measurement of reading comprehension should reflect this complexity. Kintsch and Kintsch (2005) summarized this argument well.

There is no uniform comprehension process to be measured. Instead, comprehension involves different levels and a variety of skills: The extraction of meaning from the text, the construction of the situation model, and the integration of the reader's prior knowledge and goals with the information provided by the text. These separable components of comprehension need to be assessed separately, for, as we stressed earlier, how comprehension is assessed makes a great deal of difference (pp. 86-87).

Traditional comprehension assessments were formed to optimize reliability and limit the number of distinct factors as well as the influence of prior knowledge (Kintsch & Kintsch, 2005; Pearson & Hamm, 2005). Scores from these tests include total correct or grade equivalent scores. Therefore, the information that is derived from these measures reveal very little regarding how the student organizes knowledge and forms relationships among concepts within a text to create a text representation, referred to as the situation model (Goldsmith & Johnson, 1990; Goldsmith, Johnson, & Acton, 1991; Kintsch, 1998).

Researchers have called for reading comprehension tests to conform to the following three properties. First, it is essential that the test be based in reading theory, specifically the theory of the formation of text representations (Ferstl & Kintsch, 1999; Kintsch & Kintsch, 2005). Therefore, the measure should assess a student's text representation and the complexity of this representation. Second, the measure should be easy to use, as well as provide detailed information about students' text representations (Ferstl & Kintsch, 1999). Finally, the measure should be sensitive to change over time (Ferstl & Kintsch, 1999).

The goal of the current dissertation was to create a measure of reading comprehension that would be sensitive to the complexity and multidimensionality of reading comprehension and be based in reading theory. The Passage Comprehension for Structured Text (PCST) was created as the measure to fulfill this role. Its multidimensionality was tested through the examination of the differential effects of experimental manipulations on the components of this comprehension measure. To create this measure and to determine how to test its multidimensionality, I reviewed both the psychometric literature discussing methodologies in testing multidimensionality of reading tests, as well as literature detailing reading theory and alternative reading measures.

Psychometric Studies on the Multidimensionality of Reading Comprehension

The multidimensional nature of reading comprehension has been investigated since Thorndike's first experimental studies in 1917. His key findings sparked interest in identifying underlying skills of reading comprehension that led the way for future research.

Three theories posed separate views on the number and relation of skills that underlie reading comprehension. The first theory, of independent skills, stated that reading is comprised of a multitude of independent skills and processes (Berry, 1931; Bloom et al., 1956; Davis, 1941). This theory was expanded by Davis's (1941) influential dissertation study that exhibited the use of factor analysis on reading comprehension tests. In Davis's study, as well as others following, skills were shown to factor together to form subcomponents of reading comprehension. However, others who re-examined Davis's data using slightly different methodologies found that only one factor underlies reading comprehension (Drahozal & Hanna, 1978; Rost, 1989; Thorndike, 1973). Consequently, these researchers believed that reading comprehension was comprised of only one skill and followed the theory of global skills. In this theory, reading comprehension is thought to be one global skill that cannot be separated into components or sub-skills. The third reading comprehension theory, hierarchical skills theory, was postulated when new analytical methods, such as latent trait models, were utilized (Andrich & Godfrey, 1978; Sheehan, 1997). In hierarchical skills theory, multiple, related skills are thought to underlie reading comprehension, and these skills can be ordered in a hierarchical fashion from specific to general. This theory states that reading comprehension can be thought of as one skill (the most general) and multiple skills (the most specific) simultaneously.

Although these psychometric researchers provide a variety of beliefs about the complexity of reading comprehension, it is important to note that all of the researchers in this field thought it was vital to study the multidimensionality of reading comprehension. Also, these studies provide a good account of methodologies

used to examine the multidimensionality of reading comprehension tests. In the current dissertation study, I use factor analysis to examine components of the Passage Comprehension of Structure Text (PCST) assessment. I believe the limitation of past psychometric studies is that the reading comprehension measures used in these studies were not based on reading comprehension theory.

Reading Theory

Theories of Reading Comprehension

The previous studies attempted to uncover the multidimensionality of reading comprehension through examination of the independence of skills needed to comprehend text. However, current reading comprehension theorists believe that the outcome of reading comprehension is not the ability to perform independent skills such as identifying the main idea of a passage or creating inferences, but the application of all of these skills, in concert, to form a coherent knowledge structure similar to the structure of text.

Multiple theorists have explained the process of forming this structure of text (Britton & Gulgoz, 1991, Gernsbacher, Varner, & Faust, 1990; Kintsch, 1998, Linderholm, Everson, & van den Broek, 2000). Kintsch (1998) contended that readers integrate information gained from text with background knowledge in order to summarize the content from the text. The representation that is formed from this process is a type of graphic summary of the text structure. This theory is called the “Construction-Integration Model of Reading” (Kintsch, 1998). Ideally, the final memory structure of the text created through this summarization process should capture the hierarchical nature of the text. Typically, these representations created by

readers are not identical to the text due to differences in the readers' background knowledge and reading ability, as well as the coherence of the text (Britton & Gulgoz, 1991). If the coherence of a text is extremely lacking or the reader's background knowledge is quite deficient, the resulting text representation created by the reader may be fragmented.

Two components of this text representation are formed during reading. The first component is the textbase. Textbase connections "consist(s) of those elements and relations that are directly derived from the text itself" (Kintsch, 1998, p. 103). The textbase is the underlying structure of the text intended by the author. The textbase is one component of the final representation a reader will form from reading a text.

A second component of the final text representation is the situation model. The situation model is the "construction that integrates the textbase and relevant aspects of the comprehender's knowledge" (Kintsch, 1998, p. 107). Two types of inferences form the situation model. First, one inference is used to form connections among text ideas that are not explicitly stated in the text. These inferences are used to clarify incoherent text. When a reader comes to a gap in coherence, the reader will need to hold both the textbase and prior knowledge within her short-term memory in order to find a relative bridge to fill-in the gap (Britton, Stimson, Stennett, & Gulgoz, 1998). These are referred to as "bridging inferences." A second type of inference is created to connect text information with knowledge found within the reader's long-term memory (Britton et al., 1998). These inferences help the reader to recall the text

information. Both of these inferences form the situation model. The situation model is the second component of the final representation.

While these two components of knowledge, textbase and situation model connections, form the final text representation; these connections are not represented in equal amounts. The extent to which each is present in the final representation will depend upon the explicitness of the text as well as the background knowledge of the reader. Text that is fully explicit, that demands little inference making ability by the reader would be comprised of mainly textbase connections. However, text that is unclear and incoherent would demand the reader to use more background knowledge to comprehend the text. The representation formed from incoherent text would subsequently be made of more situation model connections.

As stated previously, the goal of the current dissertation was to create a measure of reading comprehension that would be sensitive to the complexity and multidimensionality of reading comprehension. According to theory, the components of textbase connections and situation model connections are two dimensions of the final text representation formed during reading (Kintsch, 1998). The current dissertation study examined whether these components, measured with the PCST, made independent contributions to the text representation. To measure these contributions, two text manipulations were applied. I expected these text manipulations to differentially affect the two components of reading comprehension.

Coherence Effects on Text Comprehension

The effect of text coherence on recall from text has been articulated in the literature (Linderholm, Everson, & van den Broek, 2000). Experimental studies have

shown that readers are more apt to recall information after reading a coherent text compared to readers who read incoherent text (Britton & Gulgoz, 1991). Coherence of text has been improved through more detailed explication of the relationships among concepts as well as among sentences and paragraphs (Britton, Dusen, Gulgoz, & Glynn, 1989).

Findings from these studies suggest that increasing the coherence of text would facilitate text recall. However, coherence may only facilitate recall of specific components of the overall text representation. For example, Britton and Gulgoz (1991) found that coherent text helped readers form inferences among the ideas in the text; however, coherent text did not help readers recall more factual information. Furthermore, McNamara et al. (1996) showed that coherence changes in a text had a differentiating effect depending on where those changes were made. Coherence changes made to the text on the sentence level produced better recall of factual information, whereas changes made to increase the coherence of major sections within a text produced better text representations.

To increase the coherence of text, researchers have used macrosignals (Lorch & Lorch, 1996). Macrosignals are text devices used to direct the reader's attention to important information within the text. These devices include titles, headings, preview statements, summary statements, and bold and italicized print. Macrosignals help readers to discern important information and relationships present in the text. Research on the effects of macrosignals revealed that macrosignals improved the recall of information that was signaled in the text, however, the recall of information not signaled decreased. Further, additional research has found that macrosignals lead

readers to use a deeper-level strategy called the text structure strategy (Meyer et al., 1980).

Taken together, these studies lend strong empirical support to the notion that coherence changes of text will differentially affect recall of different parts of the text. In this study, I propose that macrosignals will increase the recall of the textbase component of reading comprehension more than the recall of the situation model component.

Background Knowledge Effects on Text Comprehension

The role of background knowledge on comprehension, especially its interaction with text coherence, has been studied extensively (Britton et al., 1998; McNamara et al., 1996; Pearson, Hansen, & Gordon, 1979). One question that arises out of this research is, “Do background knowledge and coherence make independent contributions to text recall or can one of these characteristics compensate for a lack in the other?” Evidence of an interaction between background knowledge and coherence has been varied and complex (McKeown, Beck, Sinatra, & Loxterman, 1992). One study revealed that text coherence increased recall of text whether or not the student had sufficient background knowledge on the text’s subject (McKeown, et al., 1992). Contrary to the previous finding, others have noted that text can be highly organized in such a way as to deter high knowledge readers from using their background knowledge to build situation model connections (McNamara et al., 1996). High knowledge readers perceived the text as easy and, consequently, did not actively process the text as needed. In that case, text coherence did not increase recall of text. Voss and Silfies (1996) presented a different view. They countered that reading

comprehension ability is necessary to build representations from coherent text, whereas background knowledge is necessary to build representations from incoherent text.

One implication of the previous studies is that coherence and background knowledge may make independent contributions to text recall. At times these two characteristics seem to act together to facilitate the process of forming a text representation, and at other times they may counteract each other's effects. In the current study, I propose that background knowledge will have a larger effect on the situation model than on the textbase, in accordance with reading theory.

Measurement of Text Representations

Since the outcome of reading comprehension is an organized knowledge representation of the text, then a reading comprehension assessment should measure this organized knowledge representation (Ferstl & Kintsch, 1999; Goldsmith & Johnson, 1990; Goldsmith, Johnson, & Acton, 1991; Kintsch & Kintsch, 2005). Traditional measures of reading comprehension that include item formats such as multiple-choice and essay questions do not provide information about a reader's text representation. More promising assessments of knowledge representations built from reading comprehension are those that use proximity data to analyze the relationships readers build from concepts within a text (Ferstl & Kintsch, 1999). Proximity data that have been collected in reading research studies have included sorting tasks (Zwaan, Langston, & Graesser, 1995; McNamara et al., 1996), cued association (Ferstl & Kintsch, 1999), and numerical rating (Bisanz, LaPorte, Vesonder, & Voss, 1978; Britton & Gulgoz, 1991). In proximity data a number is assigned to represent

the strength of relationships between select words from a reading. Words that are assumed related are given higher values than words that are not related. From these proximities a representation can be analyzed.

All of these methods for collecting proximity data have their strengths and weaknesses. The numerical rating system for the current dissertation was selected for several reasons. One important reason is that, compared to the other measuring systems, the numerical rating system provides the most detailed information about conceptual relations. A second reason is that the numerical rating system was the easiest to administer and analyze.

Pilot Investigation

To examine the multidimensionality of text, I first conducted a pilot study (see Appendix D). The pilot investigation filled two purposes. First, it tested the measurement of reading comprehension to be used in the dissertation. Secondly, it examined the independence of two components of the reading comprehension measure through experimental text manipulations. Two hypotheses were posed for the pilot investigation:

1. On the measure of textbase connections, students who read a text with macrosignals would outperform students who read a text without macrosignals, or students who read a control text.
2. On the measure of situation model connections, students who were tested on a familiar topic would outperform students who were tested on an unfamiliar topic.

Results of this initial investigation on the measure of textbase connections indicated that students who received text with macrosignals significantly outperformed both students who did not receive macrosignals and students who read a control text. However, there were no significant differences among these three groups on the measure of situation model connections.

Findings relevant to the second hypothesis revealed that students who were administered the text representation assessment on a familiar topic significantly outperformed students who were administered the assessment on an unfamiliar topic. However, there were no significant differences between these two groups on the measure of textbase connections.

These two findings suggested that the measures of textbase and situation model connections were being differentially affected by the coherence and familiarity text manipulations. Therefore, the preliminary findings suggested that these components of reading comprehension were independent.

Information gathered through an analysis of the reliability of items as well as information from topic experts gave direction on changes that should be made to the assessment. Changes were made to the texts and key words for the dissertation.

Purposes and Significance of the Dissertation Investigation

The traditional measures of reading comprehension were formed to optimize reliability and limit the number of distinct factors as well as the influence of prior knowledge (Kintsch & Kintsch, 2005; Pearson & Hamm, 2005). Although attempts have been made to update these measures to align with reading theory (National Assessment of Educational Progress, 2005), important information about student

reading is missed by current item formats. In traditional measures, reading comprehension is generally measured two ways, through multiple-choice tests or through brief or extended responses (Pearson & Hamm, 2005). Although the percentage correct and grade equivalent scores provided by these tests are often used to describe school wide and school district results, they are not informative to teachers concerning the reading abilities of individual students. The scores that are derived from these measures reveal very little regarding how the student organizes knowledge and forms relationships among concepts within a text in order to create a text representation (Goldsmith & Johnson, 1990; Goldsmith, Johnson, & Acton, 1991; Kintsch, 1998).

The use of proximity data is a promising way to measure knowledge structure built from text (Bisanz, LaPorte, Vesonder, & Voss, 1978; Britton & Gulgoz, 1991; Ferstl & Kintsch, 1999; McNamara et al., 1996; Zwaan, Langston, & Graesser, 1995). However, only a few studies have used this format with expository, structured text (Britton & Gulgoz, 1991; McNamara et al., 1996). In these studies, proximity data have not been categorized by components of reading comprehension.

An initial step of this dissertation was to create an assessment using proximity data that could measure reading comprehension of structured text. Two subtests of this measure, textbase and situation model connections, were created to assess components of reading comprehension. The first goal of this study was to determine whether these theoretical components were actually being measured by the PCST. To evaluate the PCST, the text characteristics of coherence and familiarity were manipulated. If the PCST was truly measuring the textbase and situation model

components, then the components would show sensitivity to these manipulations in a way specified by reading comprehension theory as well as past reading comprehension research. That is, coherence manipulations should positively affect textbase connections (Britton et al., 1989; Linderholm et al., 2000) and familiarity manipulations should positively affect situation model connections (Britton, et al., 1998; McNamara et al., 1996; Person, Hansen, & Gordon, 1979).

If successful, this study would extend measurement research several ways. Although proximity data have been used successfully in past studies, the measurement of components of proximity data has not been attempted. The measurement of these components would add to the usefulness of proximity data as a measure of reading comprehension. Likewise, the information obtained from the text manipulations could be used to indicate that the PCST was measuring components of reading comprehension in theoretically predictable ways.

A second goal of this study was to examine the roles of text coherence and background knowledge in forming the text representation of the reader. It has been shown extensively that both coherence and background knowledge increase text recall (Britton et al., 1989; Linderholm et al., 2000; McNamara et al., 1996; Person, Hansen, & Gordon, 1979). Theory states that text coherence will decrease the need for inferences in order to gain knowledge from text (Britton et al., 1989). Background knowledge, on the other hand, facilitates the process of forming inferences (Britton et al., 1989) to gain knowledge from text. In both cases, the inference load is decreased. A decrease in the inference load affects the number of reader characteristics that determine the amount of recall from text. Reader characteristics include the

knowledge the reader has about the text, such as amount of background knowledge or knowledge about text structure, as well as the attention the reader gives to certain aspects of the key words and the text.

When coherent text is used, the reader characteristic that will most determine how well a reader will gain text knowledge is the amount of attention the student pays to the coherent text (McNamara et al., 1996). When incoherent text is read, reader characteristics that will determine knowledge gain may involve amount of background knowledge, reading comprehension ability, as well as attention to minor text cuing. Therefore, as coherence of text decreases the number of reader characteristics that influence knowledge gain will increase. In the same vein, the number of reader characteristics that will determine situation model links changes as familiarity is increased. When text is familiar, the main reader characteristic determining how well students form situation model connections is their ability to activate their background knowledge. However, when text is unfamiliar, other reader characteristics such as the use of faulty background knowledge, attention to key word structure, as well as propensity for guessing may determine the connections readers form. Therefore, as familiarity of text decreases the number of reader characteristics determining knowledge gain will increase. The current study used factor analysis to determine the strength of the factors within the different text conditions.

Although the use of methodologies, such as factor analysis, is quite common in research exploring traditional reading comprehension assessments (Davis, 1972; Pearson & Hamm, 2005) as well as other related fields, such as reading motivation (Baker & Wigfield, 1999), few researchers have used this technique to examine

components of reading comprehension within theoretically derived, rather than skill derived reading comprehension measures. In the present study, confirmatory factor analysis was used to determine the strength of the textbase and situation model factors. The difference of fit will be measured within the different conditions formed by the text manipulations.

If successful, this study will extend theoretical knowledge in several ways. First, the differential effects of coherence and familiarity will be addressed. Information from these analyses will detail the effects of text characteristics on the structure of knowledge gained from text. Second, the use of factor analysis to test theory was also examined. If successful, factor analysis may be used to study other proximity measures of reading comprehension.

In conclusion, the role of this dissertation was to extend knowledge of educational theory and measurement research. This study was guided by two main goals. The first goal was to examine the PCST as a measurement of reading comprehension. The second goal was to examine the roles of text coherence and background knowledge in the construction of text representations. Definitions of the terms used in this dissertation study can be found in Table 1. The following hypotheses were tested in this study:

1. On the measure of textbase connections, students who read a text with macrosignals will outperform students who read a text without macrosignals.

2. On the measure of situation model connections, students who are tested on a familiar topic will outperform students who are tested on an unfamiliar topic.
3. Textbase scores will be more salient and will form a more one-dimensional factor structure when students read text with macrosignals than without macrosignals.
4. Situation model scores will be more salient and will form a more one-dimensional factor structure when students are tested on a familiar topic than an unfamiliar topic.

Table 1

Definitions of Terms

Terms	Definitional Statements
Coherence	“The continuity of meaning that enables others to make sense of a text” (NAEP, 2005).
Components of the Text Representation	Elements within the text representation. In the current study the two components are the textbase and situation model.
Episodic Text Memory	A unitary knowledge structure in the form of a network built from text and consisting of both textbase derived and knowledge derived information (Kintsch, 1994, 1998).
Familiarity	Relates to the amount of overlap between the text and the reader’s background knowledge. Highly familiar text overlaps to a great extent, whereas unfamiliar text overlaps very little.
Global Structure of Text	The local structure of the text represents the structure formed by the reader consisting of the main themes and concepts that are formed through reading multiple sentences and paragraphs Kintsch (1998).
Inference	“The act or process of deriving logical conclusions from premises known or assumed to be true; the conclusions drawn from this process” (NAEP, 2005).
Local Structure of Text	The local structure of the text represents the structure formed by the reader consisting of the information first encountered in the reading, which are the words and individual sentences (Kintsch, 1998).

Table 1 (continued)

Definitions of Terms

Terms	Definitional Statements
Macropropositions	Propositions that are formed to make up the macrostructure. These relate to information detailing the connections among paragraphs and sections of a text.
Macrosignals	Text devices used to direct the reader's attention to important information within the text. These include titles, headings, preview statements, summary statements, and bold and italicized print (Lorch & Lorch, 1996).
Macrostructure	"hierarchically ordered set of propositions representing the global structure of the text that is derived from the microstructure" (Kintsch, p. 50).
Metacognitive gap sensing	This is a sense that there is a coherence gap in the text (Britton, Stimson, Stennett, & Gulgoz, 1998).
Micropropositions	Propositions that are formed to make up the microstructure. Therefore, they relate to information detailing the sentence-by-sentence level information.
Multidimensional	Having several components.
Network	A representation of knowledge where concepts are represented by nodes and connections between the nodes are represented by links (Kintsch, 1998).

Table 1 (continued)

Definitions of Terms

Terms	Definitional Statement
Proposition	“A proposition [atomic] is simply a predicate-argument schema... (Kintsch, 1998, p.37). It represents the basic semantic unit stored in the mind of the reader.
Proximity data	Numerical scores that represent the strength of associations among concepts. The proximity is the distance between concepts within the representation and symbolizes the strength of the relationship (Britton & Gulgoz, 1991).
Salient	Prominent within the mind of the reader. Therefore, information that is salient would be the first to be recalled.
Situation Model	“A construction that integrates the textbase and relevant aspects of the comprehender’s knowledge” (Kintsch, 1998, p. 107).
Subsumption model	This model explains that during reading comprehension, information in the local structure of text becomes subsumed under main themes and ideas located in the global structure of text (Meyer, 1975).
Textbase	“The textbase consists of those elements and relations that are directly derived from the text itself “ (Kintsch, 1998, p. 103).

Chapter 2: Review of the Literature

Purpose of Literature Review

Researchers have debated the multidimensionality of reading comprehension numerous times. Although many believe reading comprehension is multidimensional, few have agreed on the number of components that form reading comprehension (Rost, 1985). However, the multidimensionality of reading comprehension and the identification of specific components of reading comprehension will have major implications on the measurement and instruction of reading. For example, if three components of reading comprehension were identified, then these three components would not only need to be used as instructional goals in a classroom, they would also need to be examined separately in comprehension measures.

Two lines of research that have examined the multidimensionality of reading comprehension will be detailed in this review. The first line of research incorporates studies that have examined the multidimensionality psychometrically by examining the number of independent components from former reading comprehension assessments. Researchers in this line have used statistical analysis such as regression analysis, factor analysis (Davis, 1941), and more currently, latent trait models (Andrich & Godfrey, 1978). These researchers have concluded that items on reading comprehension assessments have clustered around certain skills, such as identifying the main idea of a reading passage or following the structure of the text. Therefore, these researchers are concerned with the independence of the processes that form comprehension. Although these skills show the multidimensionality of reading processes, they do not attempt to relate the structure and content of the reader's

knowledge after having read a text.

The second line of research includes studies that have examined the multidimensionality of reading comprehension theoretically by explaining the components of the text structure built from reading. The emphasis of this section is on the product of reading comprehension, specifically the structure of knowledge formed by reading. These researchers, including theorists such as Kintch (1998) and Gernsbacher, Varner, and Faust (1990), explain how reading comprehension structures are formed and what information drives the construction of these structures. Therefore, these researchers are interested in the product of reading, mainly the text representation formed during reading.

Both lines of research are related and important in describing the rationale for the current dissertation. In the current dissertation, components of text representations, similar to those detailed by Kintsch (1998) were examined using psychometric techniques, such as confirmatory factor analysis and text manipulation. Therefore, understanding the literature describing the multidimensionality of the text representation formed during reading is only one step. Touching upon earlier psychometric literature relating to the multidimensionality of reading and reading skills is also necessary. After all, the skills that readers use to form their text representations will be related to the components of their text representations. I present both sets of research because I used psychometric analyses detailed in the first set of research to examine the product of knowledge that is described theoretically in the second set of research.

Psychometric Studies on the Multidimensionality of Reading Comprehension

Researchers have measured and analyzed skills important in the process of reading comprehension using many methods. Several skills were identified through experimental studies (Thorndike, 1917) and statistical analysis such as regression analysis, factor analysis (Davis, 1941), and more currently, latent trait models (Andrich & Godfrey, 1978). The goal of these studies was not only to identify important skills in comprehension, but also to explain how these skills relate to each other.

Although many studies prior to 1970 did not state theories to the relationship of skills underlying reading comprehension, theories were often implied (Chapman, 1969). These implied theories did not necessarily align with specific methods for measuring reading comprehension. Although, these theorists measured skills in comprehension quite differently, they were guided by a similar theory as to the number and relation of skills important in comprehension. Other studies used similar methods, however, they came to a different conclusion as to the importance and relation of these skills. Therefore, it is important to examine not only the methods used to measure and to identify reading comprehension skills, but the theories that guided these studies.

Examining the implied theories that guided much of the psychometric reading comprehension research has several benefits (Chapman, 1969). First, a theory would make sense of problems and concerns found in real data, such as anomalous data. Second, theories help researchers organize their thoughts and create research questions that will further the research in that area. If researchers were to find

different answers to the number of important skills in reading comprehension, if no theory is posed, the researchers may never come to an agreement or decision as to why this may have occurred. A theory will help researchers compare and contrast their findings on a meaningful level. Third, theories will guide researchers on the type of analysis that should be completed on the data. One type of analysis may be exciting and new, but researchers would have to determine if that type of analysis would provide enough information to support or reject a theory. Finally, theories allow researchers to reject, revise, or verify hypotheses.

Chapman (1969) was one of the first researchers to categorize reading comprehension studies by the theory guiding the studies, rather than the method used to measure reading comprehension skills. According to her, at least three main theories had been posed in the literature. These three theories still apply to much of the psychometric literature today and, therefore, are important to know and understand. She named these three theories as follows:

Independent Skills Theory: The independent or isolated skills theory implies that reading comprehension is a set of different processes which may be learned independently from each other and in any sequence (p. 3).

Global Skills Theory: This theory asserts that reading comprehension is a single or unitary general process, which after being learned will enable the learner to answer any kind of comprehension question about a given passage (pp. 6-7).

Hierarchical Skills Theory: The hierarchical skills theory asserts that reading skills can be arranged into levels according to the complexity of the behavior necessary to learn each skill. The acquisition of knowledge using the simpler skills is a necessary, but not sufficient, condition for the acquisition of information using the more complex skills (p. 9).

These theories, the Independent Skills Theory, the Global Skill Theory, and the Hierarchical Skills Theory, explicate the role of skills in reading comprehension. The methods and conclusions of studies exploring reading comprehension skills guided by these three theories will be reviewed.

Independent Skills Theory

The theory underlying most early studies in reading comprehension (Barton, 1930; Berry, 1931, Bloom, 1956; Gates, 1935; Gray, 1960) was that of the Independent Skills Theory. This theory implies that “reading comprehension is a set of different processes which may be learned independently from each other and in any sequence” (Chapman, 1969, p. 3). That is, the process of reading comprehension can be divided into individual processes. A reader does not just comprehend a text; instead, he will follow the structure of a passage, recall word meanings, weave ideas together, and recognize the author’s purpose (Davis, 1941). These skills are in essence sub-processes that occur separately when comprehending a text. Each skill makes a unique contribution to the process of reading comprehension. That is, each improves reading comprehension in a unique way compared to any other skill.

The Independent Skills Theory states that these sub-processes are independent

and, therefore, should be identified separately. According to this theory, as students learn to comprehend a text, the skill of summarizing a text would neither relate to nor replace the skill of finding the author's purpose. Therefore, these skills should be taught one at a time. Also, to inform a teacher on how well a student can comprehend a text the teacher needs to know the student's ability to perform these skills. Each skill should be individually tested; if a weakness is found, that skill would be practiced until the student achieved mastery of that skill (Pearson & Hamm, 2005).

Broad subjective analyses. Davis (1972) labeled very early studies guided by this theory as "broad subjective analyses." Researchers who wrote broad subjective analyses presented lists of skills believed to be involved in reading comprehension, without any empirical support to validate these lists. Many early studies were based on familiarity with the teaching of reading (Davis, 1972). Gray published an early broad subjective analysis in 1919. In the 18th Yearbook of the National Society for the Study of Education, he listed eight separate reading skills comprising comprehension:

1. To read for the purpose of giving a coherent reproduction;
2. To determine the central thought or the most important idea of a selection;
3. To select a series of closely related points and their supporting details;
4. To secure information which will aid in the solution of a problem or in answering questions;
5. To gain a clear comprehension of the essential conditions of a problem;
6. To discover new problems in regard to a topic;
7. To determine the lines of argument which support the point of view of the author; and

8. To determine the validity of statements.

In a literature review of these broad subjective analyses lists, Davis (1968) located several hundred skills. He found that many skills overlapped and could be formed into clusters of related skills. Furthermore, he realized that many simpler skills could be included in other more complex skills.

In the 1960's, researchers (Barrett, 1968; Cleland, 1965) began to form taxonomies from these simple lists. Taxonomies organized these long lists into clusters of related skills, grouping these skills into main categories and subcategories. One category, inferential comprehension, was divided into the subcategories of "inferring supporting details, inferring main ideas, inferring sequences, inferring comparisons, inferring cause and effect relationships, inferring character traits, predicting outcomes, and interpreting figurative language" (Simons, 1971, p. 343). Compared to long lists, these shortened and organized taxonomies were easier to report and study.

However, Simons (1971) argued that taxonomies also had many problems associated with them. First, the taxonomies created an unrealistic feeling of independence among skills that, in reality, had many related characteristics. For one skill, such as recalling word meanings, it is necessary to perform many different complex skills. If the skill of recalling word meanings was neatly placed under only one complex skill, the taxonomy makes the placement of this skill cleaner and more organized than is true.

Second, many skills listed were too general and not specific to reading. For example, the skill of making generalizations is used in many other problem-solving

situations. If a skill is too general, it will not be measured accurately. A better solution is to make the skill more specific to reading, such as generalizing specific information read from text into main themes.

Third, skills comprising these lists often did not distinguish between what is comprehended and how it was comprehended (Simons, 1971). For example, the skill of “making connections between new and old information” describes both the information that is being learned, and how the information is being processed through making connections. In measuring this skill, one would have to identify the “information” that should be learned and suggest an idea as to how connections should be formed. Simons stated “a description of what is comprehended is necessary before attempting to describe the comprehension process” (p. 343).

Finally, the reading comprehension skills on these lists tended to come from different stages in the comprehension processes (Simons, 1971). The first type of skill listed was the type of skill used to teach reading comprehension, such as being able to underline key words in a paragraph. Simons believed that these skills were often treated as examples of comprehension, rather than a means to forming comprehension. Just because a student may not be able to underline key words in a paragraph does not mean that he or she has not comprehended the text in a different manner. Therefore, teaching techniques, such as underlining key words, should be thought of as only one way to comprehend a text, rather than a measure of comprehension.

A second skill type, uses of comprehension, is necessary after a text is comprehended. Readers use these skills to effectively apply the knowledge gained

during reading to new situations. However, these are often not particular to reading comprehension. One skill “suggests future applications of the author’s ideas” (Simons, 1971, p. 344) is what a reader may do after she has comprehended the text. Although these skills are interesting, they should not be considered measures of the comprehension process itself.

The final skill type, comprehension processes, are actually used during the process of reading comprehension, and should be included on comprehension measures. Skills on many comprehension measures often do not add information as to how they fit into the process of reading comprehension. The description of how skills are specifically used, and when they are used during the comprehension process, needs to be explained further (Simons, 1971).

Subjective analyses with empirical support. Around the same time taxonomies of reading comprehension skills were being created, researchers were conducting studies to provide empirical support for skills underlying reading comprehension. An early empirical analyses of reading comprehension was published by Thorndike (1917). Thorndike described reading comprehension as a process of reasoning not very different from reasoning used in many other tasks, such as solving mathematical problems. He explained reading comprehension to be

a very complex procedure, involving a weighing of each of many elements in a sentence, their organization in the proper relations to one another, the selection of certain of their connotations and the rejection of others, and the cooperation of many forces to produce the final response (p. 323).

From this definition, one can note that Thorndike believed that many factors influenced reading comprehension (Pearson & Hamm, 2005). During “correct reading,” readers are able to read sentences, reject or accept new information, and organize information by making proper connections among ideas in a text. He found that poorer readers had difficulty deciding what information in a text was most important. When readers believed an idea to be much more important than its actual presentation in the text, he called this “overpotency.” When readers believed an idea to be much less important than its actual presentation in the text, he called this “underpotency.” For this reason, many readers in his studies, who had little difficulty pronouncing words in passages, had difficulty completing items on comprehension questions.

Thorndike conducted three studies to systematically examine skills involved in reading. In his investigation, Thorndike categorized the type of errors students made while answering comprehension questions, and recognized many kinds of processes used to comprehend a text. These were the first experimental studies conducted to identify what processes were performed during reading (Pearson & Hamm, 2005). Thorndike observed that although some students made no errors related to the meaning of words, these students often made errors uncovering the underlying meaning of the text. He concluded that “understanding a paragraph consists of selecting the right elements of a situation and putting them together in the right relations, and also with the right amount of weight or influence or force for each” (1917, p. 329). Thorndike’s study identified many sub-processes underlying the general process of reading comprehension. His studies have led many other

researchers to conduct studies in this domain (Davis, 1972; Pearson & Hamm, 2005).

Studies following Thorndike also examined errors made by readers while attempting to comprehend texts (Albright, 1927; Richards, 1929). Richards (1929) gave students four poems and asked his participants to comment on them. He classified comprehension errors of these poems into categories, including “failure to resist being misled by irrelevant personal memories, experiences, and associations” and “failure to discover the plain-sense literal meaning of the content” (Davis, 1972, p. 637). Albright (1927) also identified errors made by college students during reading. She formed an organized list of these 20,003 errors which included “inability to associate the related elements of the context” and “failure to grasp or retain from given explanations the ideas essential to the understanding of concepts presented later” (Davis 1972, p. 639).

Factor analysis. Chapman’s (1969) Independent Skills Theory poses the idea that sub-skills of reading comprehension can be measured independently and that these subtests of skills would be uncorrelated or only minimally correlated. Studies examining reading skills through factor analysis, on the other hand, do not believe that independent subtests could be formed that would independently measure skills of reading comprehension (Davis, 1972). Researchers using factor analysis to identify sub-skills believed that an item on a reading comprehension measure could assess more than one skill. Therefore, to isolate skills within a group of items, one would have to statistically isolate the skill. By performing principal components analysis or factor analysis, one could identify variables that explain most of the variance in a reading comprehension measure. These orthogonal components identified through

factor analysis relate to underlying skills necessary to reading comprehension.

One early study to examine these underlying skills was conducted by Davis (1941). In Davis's doctoral dissertation, he used a new statistical tool, factor analysis, to identify psychometrically the number of distinct skills that could be isolated in a standardized reading comprehension measure. He contended that comprehension could be described as a composite of nine uncorrelated mental abilities. This theory was different from Chapman's (1969) Independent Skills Theory in that many skills in a test could be correlated, but the underlying important skills, found through factor analysis, were not correlated (Davis, 1972). Davis was determined to verify that reading comprehension was a multivariate construct.

The results of Davis's study (1941) indicated that two factors, word knowledge and "reasoning about reading," were psychometrically distinct in the standardized reading test. Three other factors, although not as clear as the first two factors, could suggest that more comprehension skills may be identified in future studies. These factors included comprehension of explicit ideas, identification of text organization, and author literary devices (Pearson & Hamm, 2005). Davis noted that the reliabilities of the subtest were quite low. In following years, Davis replicated his own work with more sophisticated analyses. In 1968, his cross validation and regression analyses were used to determine the unique contribution of each skill to reading comprehension. Uniqueness was quantified by the percent of unique non-error variance of each skill. In this study, he found eight independent skills:

1. Recalling word meanings,
2. Drawing inferences about a meaning of a word from context,

3. Finding answers to questions answered explicitly or in paraphrase,
4. Weaving together ideas in the content,
5. Drawing inferences from the content,
6. Recognizing a writer's purpose, attitude, tone, and mood,
7. Identifying a writer's techniques, and
8. Following the structure of a passage (p. 541)

Davis's first findings provoked a considerable amount of attention toward using factor analysis to examine reading comprehension measures. Researchers using different reading comprehension measures and variations of factor analysis, yielded around 100 or more subcomponents, with 3-8 pertinent reading comprehension skills (Rost, 1985). The varying number of factors found in these studies was attributed to the tests used, the different groups of subjects, and the mathematical manipulations employed (Simons, 1971). This confusion as to the number of skills underlying comprehension prompted many researchers to reanalyze Davis's original data. One such study indicated that skills of word knowledge and three other comprehension skills: (1) reasoning in reading (explicit, weaving ideas, inference from content), (2) Writer's purpose, tone, mood, techniques, and (3) following structure of a passage, could be isolated psychometrically (Spearritt, 1972). However, the three comprehension skills were highly correlated, which suggested that comprehension was more univariate than previously believed. Other studies that analyzed the same data only established one general comprehension skill (Thorndike, 1973; Thurstone, 1946).

Many researchers would agree that the clearest conclusion drawn from these early factor analytic studies was that vocabulary was the most differentiated of the skills and could be considered a separate factor from reading comprehension (Spearritt, 1972). However, these early factor analysis studies failed to show how many reading comprehension skills were truly independent in reading comprehension tests (Chapman, 1969).

Limitations of Independent Skills Theory. Limitations of the Independent Skills Theory were noted in the literature. First, like the original lists and taxonomies, the skills indicated through factor analysis studies were not defined precisely (Simons, 1971). Similarly, no set labels for varying skills were assigned. Therefore, similar skills often had different names and different skills often shared the same name from study to study (Rost, 1989). Second, Davis (1968) mentioned that it would be logical to think of skills as being in a hierarchy of difficulty. He stated, “logical considerations suggests that certain of the eight skills are more basic than others” (p. 672). Davis (1941) even went so far as to partial out the influence of word knowledge from eight reading comprehension skills and found that the partial correlations were lower than the original zero-order coefficients. Although Davis (1968) mentioned that future researchers should look into a hierarchy of skills he stated that his findings “permit but do not require this situation” (p. 672).

Global Skills Theory

The Global Skills Theory “asserts that reading comprehension is a single or unitary general process, which after being learned will enable the learner to answer any kind of comprehension question about a given passage” (Chapman, 1969, pp. 6-

7). This theory of reading comprehension was formed from a reaction to the inconsistency of the number of factors identified in past factor analysis studies (Chapman, 1969). In a review of previous multivariate studies of comprehension, about 3-8 salient factors were identified (Rost, 1985). Other researchers took a slightly less multivariate approach and decided upon only two components, word knowledge and reading comprehension (Rost, 1989). All multivariate studies varied on the types of test used, participants, naming of the components, and the mathematical decisions made during the factor analysis.

This inconsistency prompted researchers to reanalyze data to determine whether reading comprehension could be considered one unitary factor, or whether comprehension was made of separate skills. Two early studies reexamined Davis's original data (Thorndike, 1973; Thurstone, 1946). Thurstone (1946) had concerns about how Davis analyzed and interpreted his results (Drahozal & Hanna, 1978). Using a slightly different methodology, Thurstone identified only one factor in the same data that Davis reported finding nine separate factors. Thorndike (1973) also reanalyzed other experimental data from Davis. He used reliability coefficients to account for error variance due to unreliability of the test. After rotation, Thorndike found three components, however, the first reading comprehension component accounted for most, 93%, of the variance in the test. The conclusion was that the Davis data was truly one-dimensional and that any variance found was due to the unreliability of the test. Thorndike assumed reading comprehension to be a form of problem solving or reasoning.

Although both Thurstone and Thorndike found one main factor accounting for the variance of the test scores, using data from Davis's original analysis examining college students, the conclusion of a general factor may have been a consequence of the age of the participants. One thought was that fluent readers, such as college students, have used skills together for so long that over time the skills become combined and could no longer be tested separately (Rost, 1989). However, beginning readers, who have not had the amount of practice and time, may not see the connections and similarity between reading skills. Beginning readers may need practice with each skill individually before using the skills together. If this was the case, one should see factors forming from reading comprehension tests for younger children. This thought led Rost to examine reading comprehension tests of Grade 2 students. German reading tests were administered to 220 second-grade students. After principal component analysis, one factor, responsible for 61% of variance, appeared as the one dominate factor in all of the items.

Researchers conducting these studies concluded that reading comprehension was a unified skill, which could not be described as a conglomerate of individual skills. The theorists went so far as to not completely dismiss the idea of reading comprehension as a measure of general intelligence evident in a verbal assessment (Rost, 1989). After all, Thorndike (1917) defined reading comprehension as a type of reasoning or problem solving not wholly different from the problem-solving ability needed when working out a mathematical problem.

Global theorists listed a few factors that could explain why researchers such as Davis found multiple components of reading comprehension. First, the analytic

methods of these studies were examined. For example, the number of participants used, content validity of variables, variable reliability, model selection, factor extraction, and factor rotation were all questioned (Rost, 1989). A second issue with multivariate studies was that researchers tended to assume different skills, if correlations were anything less than perfect. However, non-perfect correlations, according to the global theorists, could be due to any number of reasons, one being the reliability of a test (Drahozal & Hanna, 1978). One study examined the correlations of skills after correcting for attenuation. The results showed that correlations increased to a point where differentiation between skills basically disappeared (Drahozal & Hanna, 1978).

Although statistically the Global Skills Theory seemed to be evidenced in the data, the applicability of this theory to the teaching of comprehension skills has yet to be determined (Chapman, 1969). By stating that no skills underlie reading comprehension tests, teachers are given no insight into how to teach students to read. If reading comprehension is basically a problem solving skill, would teaching problem solving skills in math work just as well as having students practice reading? Further, if every reading comprehension skill were basically the same, would increasing a student's vocabulary increase his ability to find the main idea of a paragraph? According to the Global Skills Theory, this would be the case since both skills measure basically the same construct (Chapman, 1969). What if comprehension is not influenced by different skills, then would learning how to read short stories dramatically increase a student's ability to search and answer questions in a science text? Finally, if all reading comprehension skills were very similar, then teaching

more than one skill in a classroom would be a waste of time.

Hierarchical Skills Theory

For a while, researchers debated as to what skills, if any, were necessary for reading comprehension. Some researchers were quite adamant about reading comprehension being multivariate, while another group of researchers were just as adamant that reading comprehension was composed of one factor. To add to the confusion these inconclusive findings were often detected with the same data (Andrich & Godfrey, 1978; Davis, 1944; Spearritt, 1972; Thurston, 1946).

The Hierarchical Skills Theory takes into account both the strong correlation among reading comprehension skills and the belief that skills are somewhat different and distinguishable. The Hierarchical Skills Theory “asserts that reading skills can be arranged into levels according to the complexity of the behavior necessary to learn each skill. The acquisition of knowledge using the simpler skills is a necessary, but not sufficient, condition for the acquisition of information using the more complex skills” (Chapman, 1969, p. 9). According to this theory, skills can be organized along a continuum of skill complexity. Simpler skills that do not demand much complexity, such as comprehending a sentence, will be “necessary but not sufficient” for knowing more complex skills, such as making inferences that create connections among sections of a lengthy article. These skills are often arranged according to linguistic units such as words, sentences, anaphors, relationships between sentences, and inferences. While teaching a more complex skill, simple skills will often be learned. Similar to the Global Skills Theory, all skills are considered to measure one ability. Therefore, it was not surprising to find very high correlations among skills.

According to many studies, reading comprehension is comprised of at least two components, literal comprehension and inferential comprehension (Pettit & Cockriel, 1974). If a hierarchy existed, items should differentiate to one of two components with a correlation still remaining. According to the results of one study (Pettit & Cockriel, 1974), these two components did emerge, explicit understanding of text and inferred understanding of text. Each item in the measure loaded highly on one factor and also a bit on the second factor. Pettit and Cockriel determined that these data show some evidence toward a hierarchy of reading comprehension skills.

After attempting to find evidence for a hierarchical relation of skills using factor analysis, researchers decided that factor analysis was not the best way to test for a hierarchy of skills (Andrich & Godfrey, 1978). One reason was that random variation highly influenced the outcome of a factor analysis. If a uni-dimensional trait had only a small subject variance, a large number of factors would emerge. If variance in the subjects was large and traits were slightly correlated, the first factor would come out dominant. A second reason against factor analysis was that, with hierarchical data, students who mastered more complex skills in a set also mastered the lower level items as well. For this reason, correlations among skills would be quite high and the first factor would stifle all other factors. Therefore, the use of this method would be unwise when examining these different skills.

In yet another re-analysis of Davis's original data, 40 years after its collection, Andrich and Godfrey (1978) hoped to determine again, through the perspective of the Hierarchical Skill Theory, how many skills underlie reading comprehension. Andrich and Godfrey used a simple logistic model to identify which items clustered along the

one-dimensional continuum of reading comprehension. The results showed that many of the items did conform to a single dimension. Each item that fit this uni-dimensional model, the item's difficulties were estimated and graphed on that continuum. By examining this continuum, Andrich and Godfrey were able to determine a vocabulary and four comprehension subgroups of items: (a) understanding the content explicitly, (b) weaving ideas and meaning from content, (c) following the structure of content, and (d) recognizing the author's literary methodology (p. 195). It was concluded that the level of analysis determined the number of sub-skills identified in the test.

Recently, a similar analysis was conducted using a tree-based approach for assessing the number of skills in a reading comprehension measure (Sheehan, 1997, 2003). In tree-based analysis, one can determine which items have the highest discrimination at each point on a test scale. The place where an item discriminates among students is where the item is located on a scale. An easy item would discriminate among students who are lower on the achievement scale. This is because some low achievers will get the item correct and some will not. However, students who are higher on the scale will always get the item correct, therefore, the item will not discriminate among higher performing students.

This analysis takes into account both student-level diagnoses that determine individual strengths and weaknesses and group-level diagnoses that describe the typical individual profile of students who fall at a certain level on the test. A student can be described by which individual items she was able to answer correctly. A researcher could describe the strengths and weaknesses of that one student. On the

other hand, a researcher could describe what a student who is low on the achievement continuum should be able to answer correctly as well.

By determining where the items fall on this continuum, one can locate which combinations of skills are necessary at discriminating students on each level of the continuum. For instance, easier skills will discriminate students at the lower end of the continuum. Students at the upper end of the continuum, who have mastered the easy skills, will be discriminated by more difficult skills.

To determine which items had similar underlying skills Sheehan (1997) examined the Item Characteristic Curves (ICC) for each item. The ICC shows the “probability of responding correctly to items with specified skill requirements, expressed as a function of the underlying test score” (p. 335). Therefore, if a skill underlying an item is difficult, the item itself will be placed high on the continuum. If an item is made up of many easy skills, that item should fall lower on the continuum.

Items with similar skill requirements will show similar ICCs. That is, items with similar skills should be similar in difficulty. Difficulty will determine how likely that item will be answered correctly at each point on the continuum. Therefore, by examining the ICCs one could evaluate the skill mastery requirements for getting that item correct. In Sheehan (1997), eight skills were identified in the SAT comprehension test. These eight skills influenced the difficulty of the items on the test.

Hierarchical Skill Theory rectified some of the major problems associated with Independent Skills Theory and Global Skills Theory. In Hierarchical Skills Theory, skills form clusters along the reading comprehension construct continuum.

Since these skill clusters measure the same construct, it makes sense that they are very highly correlated. Second, Hierarchical Skills Theory gives some direction to teachers. Students who fall at the lower end of the continuum on a reading comprehension test should be taught lower level skills. Students who fall at the upper end of the continuum already know the lower level skills and should be directed toward practicing more complex skills. It is not the case that the skills are independent and could be learned in any order. It is also not the case that each skill is so similar that there is really no reason to teach many skills. Finally, in Hierarchical Skills Theory, skills are assumed to work together. Therefore, skills can be taught independently at first, but eventually should merge.

Theories of Reading Comprehension

Examining the psychometric studies on reading comprehension measures is important. These studies describe promising psychometric methods to assess the multidimensionality of reading comprehension measures. A limitation of these psychometric studies, however, is that the traditional measures of reading were formed to optimize reliability and limit the number of distinct factors as well as the influence of prior knowledge. Therefore, some of the variation described in these studies could be a function of the test or students themselves rather than the construct of reading comprehension. This may explain why these researchers have yet to agree on a specific number of skills or the relationship of these skills (Marshall & Glock, 1978). The following section of the paper will describe the outcome of reading posed by current reading comprehension theorists as well as alternative methods of reading comprehension measurement.

Overall, most linguistic information, including text and spoken language, has a high degree of organization and structure (Frederikson, 1977). Therefore, the knowledge acquired and stored in the memory of the reader should also have a similar organized structure. Often, this stored organization consists of “complete networks of concepts and semantic relations that connect concepts into highly organized propositional structures” (Frederikson, 1997, p. 159). During the 1970’s, psychologists began to study how readers formed this organized structure of knowledge during reading (Kintsch, 1974; Meyer, 1975). They set out to examine both the structure of an organized text itself, and the nature of how the knowledge structures stored in memory were acquired during reading. It was believed that the structure of texts could be a model to the structure of human knowledge organization (Pearson & Hamm, 2005). That is, the structure of a text often reflects the knowledge structure of the original writer (Frederiksen, 1975a). As one reads, the reader should rebuild the structure posed by the original writer of the text.

Two leading reading comprehension theories (Gernsbacher, Varner, & Faust, 1990; Kintsch, 1998) identify the reading comprehension process as a reader’s effort to build an organized, coherent structure similar to information from the text. Gernsbacher et al. referred to this process as “structure building.” According to their framework, “the goal of comprehension is to build a cohesive, mental representation or structure” (p. 431). Kintsch had a similar definition of reading comprehension and labeled the process “The Construction and Integration Model.” This outcome of reading comprehension was thought to occur automatically in good readers and was

not associated with any specific task or reading comprehension strategy. The macro-structure is the overall representation of a text that is identified by the reader.

Construction-Integration Theory

Background knowledge influence on reading comprehension. The goal of reading comprehension is to build a coherent representation similar to the structure of the original text. Kintsch (1998) referred to the process of building this representation as the Construction-Integration Model. In this model connections are formed among ideas in a text. In order to form a coherent representation, the connections that are formed have to go through a “process of integration.” The episodic text memory is built from this integration and is represented by an information network of two components.

The episodic memory is a unitary structure, but for analytic purposes it is useful to distinguish two components – the textbase and the situation model. The textbase consists of those elements and relations that are directly derived from the text itself...The reader must add nodes and establish links between nodes from his or her own knowledge and experience to make the structure coherent, to complete it, to interpret it in terms of the reader’s prior knowledge, and last but not least to integrate it with prior knowledge...The extent to which a reader will actually perform the work of transforming a textbase into a situation model is subject to numerous influences... Typically, however, the mental text representation is a mixture of text-derived and knowledge-derived information, not necessarily in equal parts.

Extreme cases, in which either the textbase or the situation model dominates to the exclusion of the other, are instructive (pp. 103-105).

The episodic memory that is constructed, therefore, is made of information from the text and background knowledge. The text information forms the textbase. Background information forms the situation model. These two components are distinct since they are built from two sources of information. However, as Kintsch (1994) described, “there exists only a single network to which both the textbase and situation model contribute links and nodes” (p. 295). Therefore, this episodic memory network is made of links and nodes from both the textbase and background knowledge.

Kintsch mentioned a hypothetical case where the reader received all of his information directly from the text. In that case, the reader would only build a textbase, not a situation model. Mostly, readers do not read a text without some knowledge of the words and information within the text. The reader brings to the text some idea of the organization of the ideas within the text.

Prior knowledge will influence the final representation built in memory (Graesser, 1998). For example, a reader may read the text *The Birth of a Baby Shark*, located in the Appendix. Two concepts within the shark text are “purse” and “cord.” Although the reader may not know how these words relate, in reference to a shark prior to reading, she may have some idea as to how a purse and a cord may be related. One reader may think of a cord as a handle to a purse. Another reader may believe that a cord and a purse have no relationship at all. After reading the shark text, the reader should understand that because of the shark egg’s shape, it is called the purse

and the cord refers to the birth cord that attaches the purse to the mother and to the newborn shark or calf.

The information presented in the text, which should help readers create new relations among concepts presented, is called the textbase. This structure does not waver from the knowledge presented in the text by the original writer. If a reader is able to integrate the information found in the textbase into a coherent memory structure, the knowledge in this structure will be easier to retrieve from long-term memory (McNamara et al., 1996).

A reader with high textbase knowledge would be able to restate a fairly accurate summary of a very explicit text. She would also be able to answer questions that relate to the explicit relationships among concepts found within the text. This student may or may not be able to elaborate further on any of the concepts or hypothesize relationships among concepts that were not explicitly related in the text.

As one reads a text, prior knowledge has some influence on how that text is read and interpreted. A reader will use prior knowledge to make sense of the information presented in a text, infer relationships among concepts within a text, or fill-in gaps of understanding while reading a text (Britton, Stimson, Stennett, & Gulgoz, 1998). Often text authors may assume the reader possesses a higher level of background knowledge than may be the case. For example, in *The Birth of a Baby Shark* a reader may need to know something about mermaids in order to determine why a shark egg's case is often called a mermaid's purse. Text authors, especially of scientific texts, often present a myriad of facts, without presenting explanations about the relationships of these facts. This can cause difficulties when there is a coherence

gap between adjacent sentences. Since the relationship between the sentences is not explicitly stated, a reader will have to infer a relationship. This inference, using background knowledge, is what creates a situation model. Often, to make this inference a reader uses knowledge in his long-term memory of information presented earlier in the text (Britton & Gulgoz, 1991). Perhaps information presented in previous paragraphs may explain the connection between two sentences. If this fails, the reader will have to use her own background knowledge to fill-in these gaps.

In *The Birth of a Baby Shark* the sentences “Baby Lemon sharks are born live” and “You can see the cord attaching the baby shark to its mother just after it was born” are found one after the other. The connection between these two facts is not made explicit. The reader would have to determine what being ‘born live’ and ‘seeing a cord’ have in common. If a reader does not see the connection, the relationship between these two pieces of information will not be formed. If, however, the reader knows something about the human baby umbilical cord, this information can be used to connection these two sentences. The ability to make these inferences while reading determines how organized the knowledge structure of the text is in the reader’s memory, and thus, determines how well the knowledge will be retrieved after reading.

Individual differences in how these inferences are formed can be found through an examination of this process. According to Britton et al. (1998), the steps in forming inferences that connect information in a text to information in prior knowledge starts first with the reader’s ability to detect one’s lack of comprehension. They termed this ability “metacognitive gap sensing,” a sense that something is

wrong with the text. In *The Birth of a Baby Shark*, the reader would have to sense that there is a gap between the sentence “Baby Lemon sharks are born live” and “You can see the cord attaching the baby shark to its mother just after it was born.”

Studies have indicated that students have varying abilities in detecting gaps in the coherence of text (Ackerman, 1986; Markman, 1979). Students with low metacognitive gap sensing will not detect incoherence and will continue to read without ever fully comprehending the text. Students high in this ability sense the gap in the text and will proceed to fill that gap with an inference, using information from another part of the text, or with prior knowledge.

Metacognitive gap sensing may not just be an ability possessed by the reader, but also a characteristic of “constraint satisfaction.” That is, a reader may choose not to fully comprehend a text if she is not constrained to do so. A reader’s purposes may set their constraint satisfaction. For example, a person reading a book for pleasure may disregard some incoherence. Yet, the same person reading to study for an exam may view the importance of comprehending differently. Therefore, the purpose for reading will dictate how well the reader pays attention to coherence gaps in the text.

If a reader determines that the text has a gap of coherence, he may be determined to rectify that gap. To rectify the gap, a reader must hold in his working memory the information preceding and following the gap in understanding, and the background knowledge associated to either of the other sources of information. In *The Birth of a Baby Shark*, the reader would have to hold in his memory the first and second sentence, and information in prior knowledge, such as information about live births and how human babies are born with umbilical cords. By simultaneously

thinking of these three sources of information, a reader may form a connection between the information to fill the gap. If this connection is well-formed, it will become part of the reader's final representation of the text, coherence will be maintained, and the reader can continue to read the rest of the text.

Three individual differences determine how well an inference is formed (Britton & Gulgoz, 1991). First, a reader needs to have a strong working memory to hold all three sources of information at one time. If the working memory is weak and the reader can only think about one or two pieces of information at a time, a connection will not be formed. Second, a reader must have specific prior knowledge activated. In the shark example, the reader must know something about live births and birth cords, either from knowledge about humans or other animals. Finally, a reader must have the ability to form an inference. A reader may be able to acknowledge the coherence gap and recall all of the information needed in working memory. However, if the reader cannot form the connection among all of the information, the text will remain incoherent to the reader.

The connections among ideas in the text and among ideas with background knowledge facilitate learning (Britton et al., 1998). As a reader forms the connection among sentences in a text, each sentence is related to the rest of the information in the text. Sentences that are not connected, either because of a gap in coherence that was ignored by the reader or not successfully filled, are left out of the text representation and are forgotten. The more sentences left out of the integration into the text representation, the less coherent and understood the whole text will be. Texts that are less understood will be more difficult to recall.

A second benefit for creating these connections in text with background knowledge is that the more connected the information is in the mind, the more likely to be activated when needed from long-term memory. The more connections a selection of information has to other information, either to other parts of text or with the reader's prior knowledge, the more likely the reader will remember that information.

Graesser and Bertus (1998) labeled connections between explicit information in text as "text-connecting" inferences and connections between text information and prior knowledge as "extra-textual" inferences. This can be seen with the shark text. If the gap is not filled, the reader will have to remember both pieces of information from both sentences separately. Using the text information and picture of the baby shark, can help the reader to remember the text information. This would be a "text-connecting" inference. However, if prior knowledge of birth (live birth/cord) is activated, the reader can also use this information to recall the information about the shark and connect to knowledge about human babies. This would be an "extra-textual" inference.

Finally, some would claim that connecting ideas and concepts in long-term memory is inherent in the processes of learning (Britton et al., 1998). Without these connections nothing is truly learned. Anything that is learned, through text or other sources of learning, forms connections between neurons in the brain. These connections among neurons will be activated when the information needs to be recalled. Without connections, information will not be able to be recalled and, thus, is not learned.

The situation model is the second component of the episodic text memory that is the outcome of text comprehension. What is referred to as the situation model is “a construction that integrates the textbase and relevant aspects of the comprehender’s knowledge” (Kintsch, 1998, p. 107). If the textbase of a reading was completely explicit, and the reader had no need to rely on background knowledge, the textbase structure and situation model structure would be the same (Kintsch, 1998). However, most texts are not fully explicit and prior knowledge is needed to make inferences and fill in gaps in a reading (Thurlow & van den Broek, 1997). Although prior knowledge may augment the representation created by the reader, it should not interfere with gaining information from the text.

The situation model will be remembered better than an isolated memory of the textbase (Kintsch, 1998). By making links from the textbase to information in long-term memory, the textbased representations will be connected to the memory in two different ways. Textbased representation will be connected as a separate “episodic memory unit,” and also to knowledge in long-term memory (McNamara et al., 1996). It will be remembered through the memory trace of the text, and through commonalities with information in long-term memory. Therefore, by forming these links, the representation formed in the reader’s mind will show a deeper understanding of text than if the reader only processed the textbase. However, researchers caution that linking prior knowledge to the representation of text will only facilitate learning if the reader has formed a coherent textbase. If the textbase is disorganized in memory, connecting information from prior knowledge may only

cause the prior knowledge to be poorly integrated into the textbased knowledge (Vidal-Ararcá & Sanjose, 1998).

The situational model represents knowledge that is situated, and, therefore, individual to the reader. In the current study, this knowledge is considered conditioned by background knowledge. Although knowledge could be conditioned by other characteristics of the text and the reader, such as task and context demands, the current study is mainly focused on the reader's background knowledge.

Local and global structures of text. The final representation is a combination of textbase information and situation model connections that form the episodic text memory. To measure this final representation created by the reader during the process of reading, one must first determine what that knowledge looks like and how it is organized in the mind of the reader. Reading theorists have stated that as a reader comprehends the text, prior knowledge that may have been disorganized should assume a structure similar to the structure of the text. Therefore, in order to infer the situation structure created by the reader, one would first have to determine the structure of the textbase. In an analysis of the textbase, one first notices that information can be presented at many different levels. First, readers take in information from individual words. After words are recognized and comprehended, readers form relations among these words to understand whole sentences. Eventually, the reader will find relationships among sentences, paragraphs, or whole chapters of a text. This level of information, from basic word-by-word to chapter-by-chapter comprehension is referred to as the local and global structure of text (Kintsch, 1998).

The local structure of the text representation refers to how the reader interprets the information first encountered in the text, which are the words and individual sentences. Kintsch (1998) referred to the local structure as the microstructure. He defined the microstructure of the situation model as the “local structure of the text, the sentence-by-sentence information, as supplemented by and integrated with long-term memory information...” (p. 50). To explicate this local structure, read the passage in Appendix A from *The Birth of a Baby Shark*. Even in the first sentence, “Some sharks can have as many as 100 babies at a time,” there is a lot of information to interpret. First, a reader needs to define key terms in the sentence such as “shark” and “babies.” If a reader does not know these words, he may be unable to read the passage or may need to use context clues. Second, a reader would need to understand all of the relations between words found within the sentence. These relations are referred to as micropropositions.

Kintsch (1998) stated that the microstructure of a text “consists of the complex propositions that comprise the text and their relationships” (pp. 65-66). Therefore, in *The Birth of a Baby Shark* one would need to comprehend individual propositions found within the first sentence. For example, what does “some sharks” mean? The reader would have to determine that the author does not mean all sharks, or infer that there are a specific number of sharks. The reader would also need to understand the proposition “can have as many as.” How many babies is the author implying that the shark can have? After comprehending, or attempting to comprehend, all of the micropropositions (propositions between words), the reader would do a similar reading of the next sentence. These simple propositions will

continue to form until the reader finds themes and main ideas that connect many of the sentences. When overriding themes of a text are identified (e.g., laying eggs and giving birth), the reader begins the process of forming the global structure of text.

The global structure of the text representation refers to how the reader interprets the main themes and concepts that are formed through reading multiple sentences and paragraphs. Kintsch (1998) referred to the global structure of a text as the macrostructure. He defined the macrostructure of the situation model as the “hierarchically ordered set of propositions representing the global structure of the text that is derived from the microstructure” (p. 50). While reading *The Birth of a Baby Shark* two themes emerge from the text: (1) some sharks lay eggs to give birth, and (2) some sharks give birth to live baby sharks. These main themes cannot be identified and processed until the reader has comprehended and related information within the individual sentences that make up the microstructure. Therefore, the macrostructure is comprised of information from the microstructure.

A second characteristic of the macrostructure is that it “organizes the propositions of the microstructure hierarchically” (Kintsch, 1998, p. 66). As the reader comprehends the details from each sentence (100 babies at a time) to the overall theme of the passage (reproduction), she may note the hierarchy that can be produced from the text. The main theme, reproduction, can be found at the top of the hierarchy; the sub-themes listed earlier can be related to this main theme, and details from each sentence can be found below the main theme. A depiction of this hierarchy in *The Birth of a Baby Shark* is located in Figure 1.

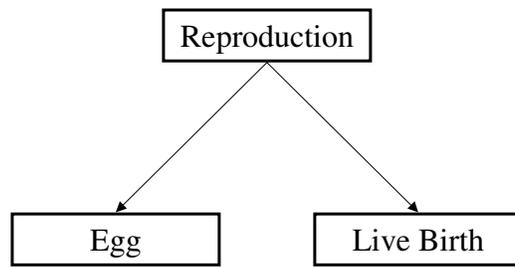


Figure 1. Hierarchical structure of *The Birth of a Baby Shark*

Other researchers have also identified the local and global structure of text through examining characteristics of text. Definitions of text characteristics listed by Marshall and Glock (1978, p. 14-17) can be arranged from local to global:

1. *Word concept*. A word concept is the underlying semantic representation of a single word (Kintsch, 1974).
2. *Relation*. A relation is a labeled and directed connection between any two word concepts (Frederiksen, 1975b).
3. *Proposition*. A proposition is the underlying semantic representation of a simple sentence as defined by case grammar (Fillmore, 1968, 1971). It is made up of at least two word concepts connected by at least one relation (Frederiksen, 1975b).
4. *Textbase*. A textbase is an ordered list of propositions. As the word concept underlies a single word and a proposition underlies a simple sentence, the textbase represents the semantic structure of a piece of connected discourse.
5. *Network*. A network is any set of labeled and directed relations. More specifically, a network can be considered as equivalent to a tree diagram in

which all the lines (relations) are labeled and indicate directionality, and in which all the spaces between relations (nodes) are identified.

The first four of these characteristics form the microstructure of the situation model as we have seen with the analysis of *The Birth of a Baby Shark*. Word concepts (some sharks) are related together to form propositions. These propositions underlie the meaning of sentences. These propositions can be listed to construct a very basic textbase of propositions from sentences. After main themes have been identified and related in a hierarchical structure, a network of ideas in a tree structure is formed. The process of building a macrostructure from the microstructure will be detailed further in the next section of this review.

The formation of the macrostructure. According to reading theorists, the outcome of reading is a coherent understanding of the macrostructure of text (Britton et al., 1998; Gernsbacher et al. 1990; Kintsch, 1998). This outcome results because a reader cannot possibly remember all the information read from a text. To facilitate the recall of the main themes and ideas from a text, a reader needs to organize and generalize the information from the text. Therefore, to comprehend a text a reader needs to build this text representation. As stated in the previous section, the macrostructure is built by organizing propositions from the microstructure into a hierarchical structure. The macrostructure can be thought of as a graphical summary of the text information. As with a summary, many detailed facts and ideas that are located in the text and processed into a microstructure are not stated explicitly in the macrostructure. Many details within a text become generalized.

One model that explains how text details become generalized is called the subsumption model. This model explains that, during reading comprehension, information in the local structure of text, from individual words and sentences, becomes subsumed under main themes and ideas located in the global structure of text (Meyer, 1975). Rather than information becoming lost, it becomes summarized into major themes.

In *The Birth of a Baby Shark*, the reader first processes individual words and phrases within sentences. In the act of comprehending the meaning of a sentence, the reader will form propositions, simple or complex relations among words and ideas in a text. Simple propositions, which form the microstructure, are called micropropositions.

Propositions formed from relating micropropositions are the base of the macrostructure. These super-propositions, called macropropositions, organize and relate the micropropositions in a hierarchy. For example, after reading the first two sentences of *The Birth of a Baby Shark* the reader will learn that sharks can lay up to 100 eggs, some baby sharks are hatched from eggs, and other sharks are born live. The reader may only grasp the ideas that some sharks are born live and others from eggs. After forming this main idea, the reader may have forgotten detailed facts such as sharks laying up to 100 eggs at a time. This “difference in birth” process main theme did not add information to the macrostructure, but instead reduced two sentences of the microstructure into a summarized form. These macropropositions superimpose the propositions formed in the microstructure. Eventually, a macroproposition will replace several micropropositions. Overall, macropropositions

are easier to remember and retrieve in long-term memory. That is why it is always more difficult to remember many disconnected facts than it is to remember one main theme.

Often in text, the macropropositions that should be formed are not explicitly stated in the text (Frederikson, 1975b). Whether or not the reader comes away from this text with a “difference in birth” process as its main theme is dependent upon the reader inferring this theme. As each new sentence is added into the macrostructure, an inference about how that sentence relates to the rest of the text is formed. The ability to form these inferences depends on the coherence of information provided in the text, the background knowledge and memory capacity of the reader, and the reader’s own standard for coherence (Linderholm, Everson, & van den Broek, 2000).

Macropropositions depend on the reader’s ability to make inferences regarding how sentences are related in the text. If the sentences do not seem to relate, bigger macropropositions stating main themes in the passage will not be formed.

Macropropositions are formed through inferences that reduce the information in a text, but will only occur if the reader understands how the information in the text can be reduced.

Eventually, macropropositions may be replaced by even greater macropropositions. In *The Birth of a Baby Shark*, multiple macropropositions were identified including differences in the birth process, birth from an egg and live birth. These propositions could be summarized into the main idea of “reproduction of a shark.” Like the process before, these macropropositions (reproduction methods) are inferred to summarize the second-order macropropositions (difference in the birth

process) previously created. Depending on the complexity of the text, many layers of macropropositions will be accumulated. Each layer forms the macrostructure of the situation model. The total macrostructure includes every layer of the macropropositions, and the hierarchical relations among these macropropositions.

Kintsch (1998) stated three rules that govern how the micropropositions are related and summarized into macropropositions. These rules were labeled “macrorules.” Selection, the first macrorule, states, “ Given a sequence of propositions, propositions that are not an interpretation condition for another proposition may be deleted” (p. 66). This rule implies that a proposition created for the microstructure will be ignored if that proposition does not fit with other propositions around it. In *The Birth of a Baby Shark* a sentence could be added to the first paragraph such as “Nurse sharks are born from eggs and can swim very fast.” The fact that nurse sharks swim very fast does not add any new information to the macroproposition of “difference in the birth process” and would not be included when forming this macroproposition. In fact, this statement would not fit into any macroproposition for this text and would be ignored when forming the macrostructure.

Generalization is the second macrorule and it states, “A proposition that is entailed by each of the sequence of propositions may be substituted for that sequence” (p. 66). This rule implies that if there is a list of similar concepts, these concepts may be generalized to a whole. Kintsch gives the example of the four chambers of the heart: (a) the right atrium, (b) the right ventricle, (c) the left atrium, and (d) the left ventricle. These can be generalized to chambers of the heart.

Construction is the third macrorule and it states, “A proposition that is entailed by the joint set of a sequence of propositions may be substituted for that sequence” (p. 66). This rule implies that at times a sequence of events can be generalized under one name. For example, a boy could select a toothbrush, place toothpaste on his toothbrush, brush his teeth, and rinse his mouth. This sequence can be placed under a more general procedure of brushing one’s teeth.

Hierarchical relations in the macrostructure. The macrostructure is comprised of multiple macropropositions formed through summarization of micropropositions. Theorists often describe macropropositions in the form of conceptual networks of important ideas in a text (Kintsch, 1998). This network, or “knowledge net,” is an associative net, with propositions in the place of nodes and links among these propositions. The links can be labeled with the strength of the association among these propositions. The meaning behind each proposition can be explained through its relations to neighboring propositions in the network.

The network of text, with directing relation among propositions, is similar to a tree diagram (Grimes, 1975) where the propositions or ideas are hierarchically organized (Meyer, 1975). Text information within the network is mutually dependent. Information from concepts and higher-order propositions are clustered into groups. These units are organized at different levels in the network. Higher-order propositions are inferentially related, although they may not be explicitly stated in the text.

The ideas within the network are grouped into the hierarchical structure based on the importance of the idea within the text. The main idea of a text is based highest in the network; details and facts are based low in the network (Meyer, 1975).

Therefore, ideas and propositions are ranked for importance. Each idea or proposition may represent one or more ideas ranked lower in importance. This ranking is often referred to as the staging of propositions (Grimes, 1975). Propositions that are “staged” high are placed at the top of the network, while low-staged propositions are found at the bottom of the network.

Different types of relations have been identified with this network of knowledge. Role relations connect propositions lower in the network, often connecting propositions from phrases and sentences. Rhetorical relations are located at the top of the network and relate sentences and paragraphs. A whole text could be considered a “very complex proposition” composed of subordinate propositions.

The macrostructure of a text is a network that summarizes the main points in a text. Kintsch (1998) described many subdivisions within a text. Concepts in the macrostructure may be directly related, indirectly related, or subordinate to another concept. Although Kintsch does not describe this relation, concepts that are not related in a text should be located further apart in the knowledge net compared to concepts that are very related in a text.

Levels of the macrostructure. Evidence suggests that upper-level propositions from a passage are recalled differently than lower propositions in a text. This relates to what Meyer (1975) and Clements (1976) referred to as the height of the information. They found that the location or height of the information within the structure of the passage predicted how well that information was retained. Information staged lower in a passage was recalled less often than information staged higher in a passage.

In examining the types of information recalled by good and poor readers, Gernsbacher, Varner, and Faust (1990) found that poor readers too often shifted their focus, and consequently formed too many substructures while reading one text. This often occurs because poorer readers cannot determine how information is related in a text. Poorer readers also give too much weight to unimportant ideas in a text. It may be beneficial to examine whether good and poor readers form the same text links when forming a situation model.

McNamara and colleagues (1996) found similar differentiation among the height of the propositions, depending on text type. In the first study, three conditions of text type were used. The first text type was the original text, which the researchers believed was locally, but not globally, coherent. They manipulated the original text to form the revised text that added “macrosignals alerting the reader to the underlying structure” (p. 6). A third text, an expanded version, was created that not only was coherent at the global level, but also was expanded to give background knowledge help to the reader.

The participants in the study, ages 11-15, were asked to read one of the texts. Following, they were directed to recall as much as they could from their reading. Their recall was scored in terms of the “proportion of propositions remembered from the text.” Propositions from each text were identified and labeled as a microproposition (staged low) or a macroproposition (staged high).

The results indicated that overall, as Meyer found in the previous study, the macropropositions were recalled significantly more than the micropropositions. However, there was also an interaction with text type. Students who read the revised

or expanded versions of the text recalled significantly more macropropositions than students who read the original version. Recall of the micropropositions was not affected by text type.

In a second study, MacNamara et al. (1996) used four text type conditions. These texts were (a) fully coherent at both the local and macrostructure levels, (b) coherent at the local level, but not the macrostructure level, (c) coherent at the macrostructure level, but not the global level, or (d) coherent for either the local or macrostructure level. As in the first study, a similar procedure for identifying propositions was used on these texts. Again, macropropositions were recalled significantly better than micropropositions. In text that included global coherence, macropropositions were recalled significantly more than micropropositions. However, for texts that were not globally coherent there was no difference in macropropositional or micropropositional recall.

Grimes's Text Representation Formation

Grimes (1975) characterized the formation of text structure differently from Kintsch, although their two theories overlap in many ways. Grimes stated that the semantic structure of text has three components. The first component of semantic structure is content, the simple textbase, which lists simple propositions found in the text. In accordance with Kintsch, this would represent a list of micropropositions used to build the microstructure. The second component is cohesion, which forms connections among these simple propositions. In relation to Kintch, these connections would be the relations among the micropropositions. The third component is staging, which refers to "the process of organizing the textbase into a hierarchy, or outline

structure, so that those propositions within the textbase that are superordinate are separated from those that are subordinate” (Marshall & Glock, 1978, p. 23). Staging is similar to building a macrostructure. According to Grimes, detailed text and ideas will be staged low in the structure and main themes and ideas will be staged high. Therefore, how a proposition is staged is determined by how important and general the concept is in the overall text.

The Structure Building Framework

Gernsbacher, Varner, and Faust (1990) described a 3-step process for how a reader forms the coherent structure of text. The first step in forming a structure is “laying a foundation.” During this step, readers try to identify the main idea of a text, often regarding the first sentence of a text as the main idea of the passage, even if the main idea is not stated until later in the paragraph or reading. The first sentence is the reader’s first clue as to the topic of the reading. So, in *The Birth of a Baby Shark* the reader may read the title and assume that the text is about shark reproduction.

The second step in the process of forming a structure is mapping incoming information. For each new sentence, a reader will determine if and how the sentence relates to the previous information. The reader may determine that the first sentence in *The Birth of a Baby Shark* relates to the reproduction of the shark. The reader will form a link between this information and what he believes is the main theme of the text. Information that is coherent and relates to previous information, such as the first sentence in the shark text, is mapped onto the same substructure as the previous information. However, if the reader believes that the new information is not related to

the previous information, as in the case where two sentences do not seem to relate, a new substructure will be formed.

The last step in the process is forming a new substructure. When a coherence gap forms in the text and the reader does not have the background knowledge to fill that gap, the information from the new sentence cannot be integrated with information from the previous sentences. When this occurs, the reader will map the new, incoming information onto a new substructure. She will see this new information as a new theme being built from the passage. However, information from the previous substructure may be harder to retrieve from long-term memory after the shift is made.

Poorer readers are less able to ignore irrelevant information. This makes them more likely to “shift” too often, to building new substructures, rather than integrating information into the structure they were developing. Their overall text structures are “bulkier, less cohesive, and less accessible” than structures formed by good comprehenders. However, less cohesive structures make it difficult to retain and retrieve information from long-term memory. Every time a shift is made to a new substructure, the information from the previous substructure becomes less accessible. Therefore, the more shifts, the less information is stored in long-term memory.

The structure-building framework has similarities and differences from the macrostructure formation proposed by Kintsch (1998). In the structure building framework a reader will first identify a main theme. Then, the reader will try to fit the incoming information with that main theme. In the macrostructure formation, a reader will build macropropositions from multiple sentences, and so forth, until coming

away with a main theme. Therefore, the macrostructure is formed from the bottom up, and the structure building framework information is organized and formed from the top-down. However, both frameworks examine how readers deal with incoherent text. That is, each sentence that is read by a reader needs to be integrated with the rest of the text. When the sentence seems unrelated to the rest of the text, the reader will either try to make the sentence relate through the use of background knowledge or through inferencing. If this does not occur, the reader may keep that information separate from the rest of the text. In either, the structure building framework or the macrostructure formation, information that does not relate to the rest of the text creates an incoherent memory structure. In both frameworks, a good reader is able to make relations among sentences and can integrate all of the information from the text into a coherent memory representation of the text. A poor reader, however, will try to remember all the information from the text without summarizing the information into a coherent memory structure.

Measuring Organized Knowledge Built from Text

Whether one believes a reader gains knowledge through a process similar to Grimes's Text Representation Formation (Grimes, 1975), The Structure Building Framework (Gernsbacher, Varner, & Faust, 1990), or Kintsch's Macrostructure Formation (Kintsch, 1998), the product of that text interaction is always some variation of an organized knowledge representation. Therefore, a measure of text interaction should also be some measure of this organized knowledge representation (Ferstl & Kintsch, 1999; Goldsmith & Johnson, 1990; Goldsmith, Johnson, & Acton, 1991; Kintsch & Kintsch, 2005). The current dissertation will focus on the

measurement of the knowledge representation from text. Although other approaches to measurement, such as portfolios attempt to measure such constructs as critical evaluation and intertextuality (Peasrons & Hamm, 2005). This study examines text representations more exclusively.

Measuring this knowledge structure using traditional assessments would be difficult. In traditional measures, knowledge is generally measured two ways, through multiple-choice tests or through brief or extended responses (Goldsmith & Johnson, 1990, Pearson & Hamm, 2005). Standardized measures of reading comprehension, such as the Nelson Denny, were formed to optimize reliability and limit the number of distinct factors as well as the influence of prior knowledge (Kintsch & Kintsch, 2005; Pearson & Hamm, 2005).

Multiple-choice reading comprehension tests often involve short passages followed by a few items requesting a reader to recall specific information from the passage. These tests could not be used as a measure of organized knowledge structures for many reasons. First, although multiple-choice test are objective and easy to score, they are quite difficult to create. The difficulty arises from the fact that direct questions are used to measure implicit knowledge. These items would be difficult to write without a substantial knowledge in both the areas of measurement and reading comprehension (Goldsmith & Johnson, 1990). Second, the multiple-choice tests cannot measure deep level understanding with such short texts and so few questions. Kintsch and Kintsch (2005) mentioned, “The texts to be read are short and performance is measured by a few questions that can be objectively scored. However, some important comprehension skills do not come into play with such short texts, and

deep understanding is not being assessed by most of the multiple-choice type questions used” (p. 87). Finally, the score on a multiple-choice test is a percentage correct. As Goldsmith and Johnson (1990) stated “a percentile ranking may be very convenient in assigning grades to students, but it tells us very little regarding what the student knows or does not know” (p. 241).

Another avenue for measuring knowledge built from text is to measure information in an essay format. On reading comprehension tests, these are generally in form of brief or extended responses. Responses to items involve answering open-ended questions about the text or basic free recall of the text. These tests, however, could not be easily used to show text-based knowledge representations. Goldsmith, Johnson, and Acton (1991) stated, “there is no simple and objective method by which the structural relations can be derived from the written answers. Because essay questions permit a wide variety of responses, it would be extremely difficult to derive the structural relations for the complete set of concepts” (p. 88). Essay type assessments, like multiple-choice tests, use percentage correct as a score on the test. This percentage correct, however, cannot show how a reader has organized concepts within her text representation.

Latent semantic analysis (LSA) is a new methodology for scoring essay questions in order to assess the underlying structure. Kintsch and Kintsch (2005) describe LSA as:

A machine learning method that constructs a geometric representation of meaning that resembles the structure of human knowledge about

words and texts. It constructs this representation simply from observing how words are used in a large number of texts.

Through this method, the representations of text as well as essay statements of text can be represented and even compared. Therefore, this method could be used to score essay tests.

A second non-traditional way to measure text-based knowledge is through proximity data. Proximity data are numerical scores that represent the strength of associations among concepts. The proximity is the distance between concepts within the representation and symbolizes the strength of the relationship. Although this technique is not new in the reading comprehension field, few studies have applied it (Ferstl & Kintsch, 1999). These methods include sorting tasks, cued association, and numerical rating (Ferstl & Kintsch, 1999; Goldsmith, Johnson, & Acton, 1991). All of these methods result in some form of proximity data, which can be used to measure knowledge representations.

One way to collect proximity data is through a sorting task. In a sorting task, participants are asked to sort concepts into a pre-described number of categories. A sorting task was implemented in a Zwaan, Langston, and Graesser (1995) study. The study's undergraduate subjects were asked to read a narrative. Ten words (verbs) from the narrative were listed on a following page along with seven empty columns. The subjects were asked to organize the 10 words by writing them into the seven empty boxes, as they choose. Proximity scores were generated for each pair of words from the data. The proximity score represented the percent of students who had sorted two (a pair) of the 10 words into the same category. The scores ranged from zero to

one. A score of zero was given if no students placed a pair of words into the same list. A score of one was given if all students placed the words into the same list. This task was completed for three more narratives. The 10 words selected in each story were analysed by their textbase (distance in the text and argument overlap) as well as by their situation representation (spatial relatedness, protagonist identity, causal relatedness, and intentional relatedness). A vocabulary measure was also included. When students were unable to see the text while making their sortings, the five situation variables as well as the vocabulary test predicted the relatedness of the verbs. When students were able to view the text while making their sortings, the five situation variables, the vocabulary measure, and the surface distance of verbs were significantly related to the relatedness of the verbs. This study indicated that the situation model, the textbase, and a vocabulary measure each make contributions in the relatedness ratings of verbs in a fictional text.

A sorting task was also used to measure reading comprehension in McNamara et al. (1996). In this task, students were asked to sort 16 animal characteristics into as many piles as they saw fit. Seven of these characteristics were pre-classified as mammalian, four were classified as mammalian or non-mammalian, and five were classified as strictly non-mammalian. These three categories were considered the optimal sorting for the characteristics. After reading a text about animals, the students were asked to perform this sorting for a second time. A similarity matrix was created to summarize how many times a characteristic was sorted into the same category with each of the other 15 characteristics in the pre and post-test. This number was the proximity score and indicated the strength of the relationship between concepts. The

more times a characteristic was sorted with another characteristic the higher the similarity. Also, individual scores were given to suggest how close each student was to the optimal sorting. The results of this study showed that the second sort produced more piles related around mammalian characteristics than the first sort, showing that the students' changed their representation of knowledge after reading the text. The data collection by the sorting task is fairly simple, and this method can measure text representations with many different concepts. However, the information obtained from this method may not create the best indicator of relations among the concepts. Ferstl and Kintsch (1999) stated that in a sorting task "the resulting structures for each individual are relatively sparse. All relationships are assumed to have equal strength, and items in the same group are not differentiated at all" (p. 250).

A second proximity data collection technique is through cued association. Ferstl and Kintsch (1999) used a cued association task to measure reading comprehension of undergraduate students on a narrative about a birthday. To create this assessment the researchers first identified 30 concepts that related to birthdays in general. These were collected through a free-association task with a group of participants. Next, the researchers selected 30 key concepts from the birthday narrative. A prepositional structure of the narrative was used to determine the association between the 60 words on the list. During the assessment the students were shown all 60 words. Afterwards, the students were given each word and asked to name one to three words that first came to mind. From these words, proximity data was created. The first word stated in the association list was considered to have the strongest association with the cued word, the second word having the second

strongest association, and the third word the third strongest relation. Although this method seemed to provide a rich source of information, it was quite time consuming to select both key concepts from a reading, choose prior knowledge items, and score each individual test. If the data collection and scoring could be computerized to make it less time consuming, this method would be promising.

Lastly, numerical ratings have been used to measure reading comprehension. One of the first studies to use numerical ratings for reading research was Bisanz, LaPorte, Vesonder, and Voss (1978). Bisanz, et al. asked undergraduate students to rate the similarity of nine animals, two at a time on a scale from 1, high similarity, to 7 low similarity. This represented a prior knowledge measure. Next, the students were asked to read a story about the nine animals. Following the reading, the students were asked to rate the similarity of the animals again, keeping in mind how the animals were described in the story. By using MDS, a visual depiction of the scores was viewed. The prior knowledge visual depiction of the relations among the animals was different than the post reading visual depiction. Therefore, reading the text changed the similarity ratings the students gave the animals. MDS is used with aggregated data; therefore, individual scores could not be obtained and examined. Although these early data were quite interesting in that they showed a direction for measuring text representations, they were not based in a text comprehension theory (Ferstl & Kintsch, 1999). A more recent study by Britton and Gulgoz (1991) used numerical ratings. Britton and Gulgoz selected 12 key terms from one of three readings. One reading was the original text and the other two readings were increased through coherence changes. Air Force recruits were asked to rate the relatedness of these 12

words from the readings on a 7-point scale, two at a time, for a total of 66 ratings. Seven experts and the text author were also administered the test. The recruits' ratings were correlated with the experts' ratings as a measure of reading comprehension. The recruits in the coherence text treatments outperformed the recruits in the original text treatment.

In the Britton and Gulgoz (1991) study, a computer program called Pathfinder was used to collect and analyze the data. The data collection program in Pathfinder randomly displays the pairs of words and asks for the test taker to rate the words using the keyboard. Britton and Gulgoz (1991) randomly showed participants 66 pairs of words. After the student finished rating all pairs of words, the computer program stored that student's information in a computer file. Pathfinder also has a data analysis program. This program provides correlation scores, coherence indexes, as well as a visual depiction of the representation using an algorithm that finds the best "path" among the variables. Pathfinder has been used to examine knowledge from other domains such as pedagogical knowledge (Housner, Gomez, & Griffey, 1993), psychology (Gonzalvo, Canas, & Bajo, 1994), biology (Cooke, 1992), and statistics (Goldsmith, Johnson, & Acton, 1991).

In the current dissertation, the proximity data was collected through numerical ratings. This method was selected for several reasons. First, numerical ratings provide "rich knowledge structures specifying a relationship for each pair of concepts" (Ferstl & Kintsch, 1999). Second, the similarity ratings were easy to collect and to use to assess knowledge structures. Third, Britton and Gulgoz (1991) used numeric ratings for one of the few research papers to use proximity data with nonfiction text.

Therefore, since this was previously successfully method, I believed that it would work in a similar situation. Finally, since a relationship or lack of relationship is specified for every pair of concepts, these relationships could be compared among all the students.

Variability of Text: Coherence and Background Knowledge

Variability in texts can influence how a reader gains information from a text to form a text representation (RAND Reading Study Group, 2002). Two major text variations will be examined in the present study. The first variation is the content covered in a text. This includes the “age-appropriate selection of subject matter,” and text familiarity (p. 25). A second variation is the coherence and readability of a text. The RAND group stated that it is important to conduct research on text variability in order to design “effective instruction and informative assessments” (p. 26). Therefore, the current study will examine the extent that a reading assessment is sensitive to these text variations.

Research on Increasing Coherence of Text

Many studies have examined the effects of text coherence on a reader’s ability to build a coherent memory structure of text (Linderholm et al., 2000). The number of gaps in the textbase determines the coherence of text. These gaps make establishing relations among details difficult for the reader. Previously, it was noted that when a reader encounters a gap in the textbase, he first determines if information earlier in the text can be used to make sense of the coherence gap. If that strategy fails, a reader often retrieves prior knowledge to help make sense of the coherence gap. However, if

a reader does not have the prior knowledge needed, the structure the reader creates from reading the text will be incoherent and difficult to retrieve from long-term memory.

Therefore, studies have been conducted to make text more coherent so that readers will not have to actively search for inferences to relate the incoherent information (Britton et al., 1989; Linderholm et al., 2000). By making connections among ideas more explicit, students will form a more coherent macrostructure. By facilitating this process, readers need to make fewer inferences during reading. Linderholm and colleagues believed that the best way to facilitate inference generation was by improving the causal text connections. In their study, Linderholm et al. made sure that the temporal order of the information in the text was in the correct direction. That is, the “cause precede[d] its consequent in the text” (p. 532). If the order is reversed, readers will use more energy to rearrange the events to determine what is causing an event. Second, Linderholm and her colleagues made sure goals in the text were explicit, such as the main themes. This made it easier for readers to form the hierarchical macrostructure. Third, coherence breaks were corrected. This occurred when “inadequate explanations were provided by the text, multiple causality was involved, or the current sentence and its explanations were separated by large segments of text” (p. 533).

Results indicated that for both less-skilled and more-skilled college readers, the revisions to the text improved comprehension and recall. Actually, when the text was revised, less-skilled readers performed as well as the more-skilled readers. This study showed that the process of forming a coherent macrostructure relies on

inference-generation. Inference-generation is difficult to do and takes quite a bit of prior knowledge and memory resources. When fewer inferences need to be made from a text, the process of comprehending the text becomes easier.

Britton and Gulgoz (1991) used the Construction-Integration Model to make referential repairs to text. In this study, they identified places in the text where sentences were not referred to by a previous sentence. These breaks in coherence made it more difficult for the reader to build connections between sentences. To improve the text, they made references to all concepts explicit, used the same terminology for major concepts throughout the text, and arranged sentences so that the beginning of the sentence showed how the sentence fit in with the rest of the text. The results revealed that more information was recalled and more inference questions were answered correctly when college students read the repaired text. Correctness on items that asked for factual information did not differ between the original text and the repaired text. Britton and Gulgoz's theory was that forming a coherent, mental representation with links among major concepts in a text would not help a reader recall specific details in a text. This would be similar to the macro-structure formation. That is, when the structure is formed, details and instances become generalized. Having a coherent macrostructure would not help a reader retain detailed information.

In their next study, Britton and Gulgoz (1991) examined the actual shape of the representations formed during reading and compared these to expert models. In this measure, they selected 12 terms from a text and presented these terms two at a time for a total of 66 pairs of words. After reading, Air Force recruits were asked to

rate the relatedness of the terms on a seven-point scale. Seven experts and the author of the text were also administered the test on the original text. Each recruit's ratings were correlated with each expert's ratings. The correlation score indicated how close the recruits' mental models were to the experts' models. Higher correlation scores were evidenced from the recruits who received the revised text compared to the recruits who received the original text. Therefore, the revised texts helped the recruits form a structure similar to other experts and to the intentions of the author of the text.

Increasing Coherence Through use of Macrosignals

Macrosignals are used to create coherence within a text by guiding the reader to form macropropositions (Loman & Mayer, 1983; Lorch & Lorch, 1996; Rickards, Fajen, Sullivan, & Gillespie, 1997). Macrosignals are "writing devices that emphasize aspects of a text's content or structure, without adding to the content of the text" (Lorch, 1989, p. 209). Types of macrosignals include titles, headings, previews, summaries, typographical cues, and number signals (Lorch, 1989; Meyer, 1975). The effect shared by all signals is to direct the attention of the reader to important information and the hierarchical organization of expository text (Loman & Mayer, 1983; Lorch, 1989; Lorch & Lorch, 1996; Rickards et al., 1997; Spyridakis & Standal, 1987). However, this attention is directed at different content and sections of the text representation, and different amounts of the text depending on the type of signal used (Lorch, 1989). The effects of three select macrosignals will be discussed and manipulated in the current study.

Headings are one important macrosignal that label the sections and subsections within a text (Brooks, Dansereau, Spurlin, & Holley, 1983; Dee-Lucas &

Di Vesta, 1980; Lorch, 1989; Lorch, Lorch, & Inman, 1993; Mayer, Dyck, & Cook, 1984; Spyridakis & Standal, 1987). Headings act as hierarchical outlines embedded within the text (Brooks et al., 1983). The scope of the information signaled by the heading depends on the amount of information following the heading and its overall fit with that information. Through this labeling, the reader will build a macrostructure more similar to the macrostructure intended by the author than a reader who is not given headings (Lorch, 1989). This occurs for many reasons. First, headings serve to activate the background knowledge of a reader (Brooks et al., 1983). The reader is able to use her knowledge to make connections within the text. Second, headings highlight what is important within the text of a particular section (Lorch, 1989). The reader is able to select what should be remembered. Third, the heading acts as a retrieval cue during recall (Brooks et al., 1983). Headings help recreate the text representation that the reader can use as a memory tool while attempting to recall the text. Since the information is more organized within the memory of the reader, headings should at least increase memory for the information cued by the headings if not for the entire text (Lorch, 1989). Fourth, headings highlight relationships among the superordinate and subordinate information within the text (Brooks et al., 1983).

Both overviews and preview statements have been used as well to increase the reader's memory for text (Lorch, 1989; Lorch & Lorch, 1985, 1995; Lorch, Lorch, & Inman, 1993; Mayer, Dyck, & Cook, 1984; Rickards et al., 1997; Spyridakis & Standal, 1987). Whereas overviews are used like abstracts in a journal article to relate the information from the entire text, previews often describe information in one section or paragraph (Lorch, 1989). Like headings, overviews and preview statements

state the content of the text before the reader gets into the body of the work.

Therefore, like headings, overviews and preview statements can activate background knowledge and highlight important information from the text. However, overviews, which cover a larger span of information, act as an outline for the entire text.

Therefore, overviews not only give information about the major, superordinate concepts in the body of the text, but they also provide information about their interrelatedness (Brooks et al., 1983). Overviews influence the number of main concepts recalled after reading, but not the recall amount of subordinate information presented about each main concept (Lorch, 1989). Headings and preview statements only cover the span of one section, and therefore aid in the construction of superordinate and subordinate relationships (Brooks et al., 1983).

Finally, typographical cues have been used to cue very specific information in a text (Cashen & Leicht, 1970; Crouse & Idstein, 1972; Frase & Schwartz, 1979; Lorch, 1989; Rickards & August, 1975). These cues include bold-faced words, italics, underlining, capitalization, font color, font style, and different spatial placement. The main job of typographical cues is to direct attention at very specific words, phrases, or sentences. At times, other signals, such as headings and titles, can also have a typographical cue, such as a larger font and being set off from the body of the text. Typographical cues can also be used to highlight key words or topic sentences within the body of the text. Compared to headings, overviews, and previews, typographical cues influence the recall of a very small amount of information from a text. The impact of a typographical cue on recall is related to the amount of typographical cues used in the text. The less the typographical cue is used, the more distinctive the

information that is cued, and the more influence it will have on retention (Lorch, 1989).

Studies have shown that the influence of any signal, whether it is headings, overviews, or typographical cues, depends on the complexity of the text (Lorch & Lorch, 1996; Spyridakis & Standal, 1987). For example, Lorch and Lorch (1996) examined the effects of the combination of topical overviews, headings, and topical summaries on the recall of text with simple and complex topic structures. The researchers measured memory of the text with tests measuring the reader's overall recall, the number of topics recalled, the proportion of the ideas recalled in relation to a topic, and the order of the topics recalled. Overall, signaling influenced the recall of the organization of the topics and the amount of subordinate information recalled under each topic. The results indicated that these effects were greater for those reading the complex topic structure rather than the simple topic structure. The authors note that other researchers, who did not find significant results for signals, may not have been using text that was complex enough.

Another reason why studies may not find a significant increase in recall for signaled text is that while signals increase memory for the signaled text, they actually inhibit the memory of text that is not signaled (Lorch & Lorch, 1996). For instance, Lorch and Lorch had subjects read a text where half of the text was signaled with headings and half was not signaled. Subjects who received the half-signaled text recalled the signaled information more than the students who received the no-signals text. However, the subjects who received the half-signaled text did poorer recalling the information that was not signaled than subjects who read the no-signals text.

Therefore, the presence of signals interfered with the subjects learning the information that was not signaled. The authors state that readers believe that non-signaled text, when signals are used elsewhere in the text, is not important to remember. These readers put a greater effort in learning the signaled text. Therefore, if one were to measure overall recall, there would be no difference between subjects with signaled text versus subjects without signals.

Signals not only change the attention readers give to the information in the text, they influence the type of reading strategies used to read the text (Mayer, Dyck, & Cook, 1984; Meyer, 1980). Mayer, Dyck, and Cook believed that readers who were given signals in the text were more likely to use a “mental model strategy” than readers without signals. In the mental model strategy, readers created a coherent representation of text from the signaled information. This was shown by the increased recall for important signaled information in the text and students’ increased ability to answer problem-solving items. Readers without signals were more likely to use a list-strategy to memorize information from the text. In this strategy, the reader remembered the information presented in the text as a “meaningless” list of information. This was shown by the increased recall for information from the beginning of the passage more than from the rest of the passage compared to readers with signals. In their study, students received text with or without signals, including preview statements and headings. Recall was tested with a recall test, problem-solving questions, linking questions, and verbatim recognition items. Students with signaled text outperformed students with the unsignaled text on problem-solving and linking, but not verbatim recognition. Therefore, these readers were using a mental

model strategy to form a coherent mental representation of the text to solve problems and form links among information in the passage. After review of the recall test, they found that the subjects with the unsigaled text actually remembered more from the beginning of the passage than the students with the sigaled text. This led the researchers to suspect that the students used a list strategy rather than a mental model strategy.

Reader characteristics can also determine a student's strategy use when encountering a sigaled text (Meyer, Brandt, & Bluth, 1980). Ninth-grade students with varying levels of comprehension and vocabulary knowledge took recall tests after reading text with or without signals. Meyer et al. (1980) examined the organization of the information found in the recall essays to determine if these were similar to the author's organization. Results indicated that students with high comprehension formed macropropositions automatically, that is, they organized their writing similar to the author, with or without signals in the text. Students with low comprehension and low vocabulary scores used the list strategy automatically, with or without signals in the text. However, students with lower comprehension scores than vocabulary scores formed macropropositions only when signals were added to the text. A list strategy was used when no signals were used.

Increasing Coherence Through use of Text Structure Strategies

Another way to help readers identify important information in a text and to create coherent text representations is by teaching readers to use text structure strategies. Text structure strategies help the reader follow the hierarchical organization of the text. Readers who are taught how to use the structure to extract the

main ideas from a text remember more of the text and recall more important information than readers who do not use this strategy (Meyer et al., 1980).

In text-structure training, readers “learned to identify and use basic, top-level structures to organize their ideas... learned to recognize these structures in everyday reading materials, and to use these structures as a framework for acquiring new information” (Meyer & Poon, p. 146). To identify the structures readers are trained to use text signaling. Following are five, basic, top-level hierarchical structures of texts (Meyer, Middlemiss, Brezinski, McDougall, & Bartlett, 2002, p. 490):

1. **Description:** Descriptive ideas that give attributes, specifics, or setting information about a topic. The main idea is that attributes of a topic are discussed (e.g., newspaper article describing who, what, where, when, and how).
2. **Sequence:** Ideas grouped on the basis of order or time. The main idea is the procedure or is history-related (e.g., recipe procedures, history of Civil War battles, growth from birth to 12 months).
3. **Causation:** Presents causal or cause-and-effect-like relations between ideas. The main idea is organized into cause-and-effect parts (e.g., directions: *if* you want to take good pictures, *then* you must....; explanations: the idea explained is the effect, and the explanation is its cause).
4. **Problem-solution:** The main ideas are organized into two parts: a *problem* part and a *solution* part that respond to the problem by trying to eliminate it, or a *question* part and an *answer* part that respond to the question by trying to

answer it (e.g., scientific articles often first raise a question or problem and then seek to give an answer or solution).

5. Comparison: Relates ideas on the basis of differences and similarities. The main idea is organized in parts that provide a comparison, contrast, or alternative perspective on a topic (e.g., political speeches, particularly where one view is clearly favored over the other).

In their study, Meyer and Poon (2001) compared the structure-strategy training with the interest-list strategy training, and a control condition. Participants in the interest-list strategy training were encouraged to evaluate texts for interestingness and to practice remembering the information in the text. If the participant found the text to be uninteresting, she was encouraged to think of a friend or relative who would find the article interesting and share the information with that individual. Participants in the control condition received no training.

The researchers tested the participants on many measures. The number of propositions recalled from the text measured total recall. The “gist” was measured with multiple summary tasks, including questions about the main topics in the text. A top-level structure measure assessed the order of the propositions in the recall of the text and compared them to the order of the propositions following the structure of the text. Compared to the other two conditions, participants in the structure strategy training condition recalled significantly more information from the text, produced better summaries, answered the summary questions better, and produced greater top-level scores.

Overall, this study showed that training students to form a macrostructure from a text increased their memory for the important information of the text. This aligns well with Kintsch's macrostructure theory. Readers usually cannot remember all the information in a text. However, if readers use the macrostructure of the textbase to form a situation model from this macrostructure, information should be easier to encode into memory and easier to recall.

Research on the Influence of Background Knowledge

It has been widely documented that prior knowledge increases recall of text (Britton et al., 1998; McNamara et al., 1996; Pearson, Hansen, & Gordon, 1979). Recently, researchers have been interested in not only examining this effect, but how this effect changes depending on the characteristics of the text and the learner. Text coherence is one text characteristic that has been shown to interact with background knowledge.

Two possible interactions between background knowledge and text coherence may occur (McKeown et al., 1992). First, prior knowledge may fully compensate for lack of coherence in a text. If a student were to read a text with familiar information, lack of coherence in that text would not be an issue. Second, having both high background knowledge and coherent text may produce better recall than only having one characteristic.

To test the effect of high knowledge on comprehension of text, McKeown et al. (1992) enhanced student background knowledge by presenting information prior to reading that would make the content more familiar to the student. Next, the students were randomly assigned to read two texts with varying levels of coherence.

Students then were assigned to read an original social studies text or a revised text. Text revisions were made to facilitate students forming coherent representations and to make links with prior knowledge. These revisions were also intended to “establish textual coherence by clarifying, elaborating, explaining, and motivating important information, and making relationships explicit” (p. 80). Reading comprehension was measured with open-ended questions.

Results showed that even with the background knowledge exposure, students who read the revised test recalled more information than students who read the original version. This study demonstrated that coherence of text is needed, even when readers are somewhat familiar with a topic. McKeown et al. (1992) concluded, “background knowledge is most useful if the text is coherent enough to allow the reader to see the connections between the text information and previous knowledge so that the knowledge can be combined with the text information to create a meaningful representation” (p. 91).

McKeown et al. (1992) found that for textbased recall, knowledge did not affect the influence of coherence. However, how would knowledge affect the influence of coherence if one were to measure higher-order questions that involved deep understanding of a text and the formation of a situation model? McNamara et al. (1996) attempted to answer this question.

In McNamara et al. study (1996), four different texts were formed by manipulating the coherence levels of a text at both the microstructure and macrostructure levels. These included texts that were (a) fully coherent at both the microstructure and macrostructure levels, (b) coherent at the microstructure level, but

not the macrostructure level, (c) coherent at the macrostructure level, but not the microstructure level, and (d) incoherent at either the microstructure or macrostructure level.

Comprehension measures were administered to evaluate a reader's recall of the textbase and the reader's ability to form a situation model. Free recall measured a reader's recall of the textbase. Textbased questions measured a reader's ability to recall information from single sentences. Measures that examined a reader's ability to form a situation model included problem-solving questions, bridging-inference questions, elaborative-inference questions, and a sorting task. Problem-solving questions measured a reader's ability to apply information to novel situations. Bridging-inference questions measured a reader's ability to form relations among sentences, while elaborative inference questions measured a reader's ability to form connections between textbase information and the reader's prior knowledge. Finally, the sorting task measured "understanding of the relation between concepts presented in the text" (p. 19). This task required the reader to sort main concepts into as many piles as needed.

The effect of prior knowledge, measured with multiple knowledge tasks, was examined in relation to the four text conditions, the measures of textbase recall, and situation model formation. Overall, the students recalled the least coherent text most poorly of all the text conditions on the recall measure. Also, readers with greater prior knowledge outperformed lower prior knowledge readers on the posttest questions. However, by examining the posttest questions by knowledge, an interesting pattern emerged. Low-knowledge readers performed best on all types of questions and the

sorting task when presented with coherent texts, and as well as the high-knowledge readers on the textbased questions. However, on the three situation-model questions and the sorting task, the high prior knowledge readers performed best when the fully incoherent text was presented.

McNamara et al. (1996) explained how their results supported the theory of building a situation model and macrostructure. They believed that “active processing,” that is, building coherence by relying on memory capacity and background knowledge, was preferable to students with high background knowledge than reading a coherent text that did this work for them. In incoherent texts, readers pull knowledge directly from their own knowledge store and use it to rectify the coherence gaps. By doing this, readers are making more ties from their personal, prior knowledge to the textbase they have built. Therefore, knowledge is more integrated into the mind of the reader, making the text understood more deeply and easier to recall. A student with low prior knowledge does not have this source of knowledge, therefore, gaps in coherence are never filled and the text is never understood.

It seems then that when text is less coherent, the reader makes more use of background knowledge to make sense of the text structure. In the McNamara et al. (1996) study, high-knowledge readers performed best when given low-coherence text. The low-coherence text forced the reader to pay more attention to the text, and to integrate more background knowledge into his text representation. Therefore, one may wonder, to comprehend a text, when does a reader use comprehension strategies and when does he use background knowledge?

Voss and Silfies (1996) hypothesized that text-explicit connections among ideas made the use of background knowledge unnecessary. For example, the high-knowledge readers in the McNamara et al. (1996) study integrated background knowledge and formed a situation model only when the text was minimally coherent. These students were forced to use their background knowledge to make sense of the reading. However, when text is coherent, prior knowledge is not essential, and it is more important for the reader to have high comprehension skills to extract the appropriate text representation.

However, when text is incoherent, prior knowledge is needed, as reading comprehension skills do not fully help the reader to form a coherent structure. Therefore, reading comprehension skills are needed in forming a representation in a coherent text and prior knowledge is key in forming a representation in an incoherent text.

To show this relationship of prior knowledge and text coherence, Voss and Silfies (1996) measured both reading comprehension, using the Nelson Denny, and background knowledge prior to reading. Students were assigned to read a coherent or incoherent text. Following, the students answered 10 recall questions. In the coherent text group, only reading comprehension skill related significantly to recall and in the incoherent text group, only background knowledge related significantly to recall. Voss and Silfies (1996) explained that when text is coherent, the reader is more likely to build a textbased representation that requires a high degree of comprehension skill. However, when the text is incoherent, the reader is more likely to build a situation model that requires a high degree of background knowledge.

In the previous studies, the level of background knowledge was a characteristic of the reader and not the text. In McKeown et al. (1992) the readers were taught the background knowledge before reading. In the studies by McNamara et al. (1996) and Voss and Silfies (1996), readers were tested for their level of background knowledge. Vidal-Abarca and Sanjose (1998) approached the background knowledge variable differently by changing the text rather than the reader. To change the text, they not only increased its coherence, but they also increased linkages to background knowledge. Coherence was enhanced through additions of headings, summaries, and surface structures. Links to background knowledge were augmented by the use of familiar imagery and additional information to aid understanding of important relationships and conclusions in the text. The study had four text conditions: (a) the original text, (b) the coherent text, (c) the linking text, and (d) the coherent and linking text.

After reading a text, students were asked to recall as much of the passage as possible. Overall, recall was best for students who received the coherent and linking text. In further analysis, students who read the coherent text recalled more statements related to the text than the students who read the linking text. Students who read the linking text produced more statements that included additional information not found in the text than students who read the coherent text.

The overall results indicated that both coherence and linking to background knowledge increased the recall of high-level units over the original text. This verifies the situation model. The situation model, according to Kintsch (1998) is a combination of a coherent textbase understanding and connection to prior knowledge.

Therefore, the finding of this study, that both links are important for increasing text comprehension of higher-order macropropositions, aligns well to the formation of the situation model.

Theoretical Expectations and Hypotheses

The purpose of this dissertation was to examine the multidimensional nature of reading comprehension. I proposed that there are two independent components of reading comprehension. These components, the textbase and situation model, have been described in Kintsch (1994, 1998). Past studies have indicated that coherence and familiarity differentially affects these two components (McNamara et al., 1996). In this dissertation, coherence and familiarity were manipulated and the differential effects on textbase and situation model connections within a text representation were explored.

Expectations

Coherence. The first step toward understanding the independence of these two text representation components is to document the effects of text coherence on the ability of readers to create a coherent textbase and situation model. The positive effect of coherence on comprehension and text recall has been found in numerous studies (Britton & Gulgoz, 1991; Linderholm et al., 2000; Lorch & Lorch, 1996). Recall is significantly increased by reading coherent text compared to incoherent text (Britton & Gulgoz, 1991). Coherence of text can be manipulated by adding more explication of the connections among text concepts and by improving flow among sentences and paragraphs (Britton et al., 1989). Macrosignals (titles, headings, statements, summary

statements, and bold and italicized print) have also been used to increase text coherence (Lorch & Lorch, 1996). These signals direct a reader's attention to important relationships in the text.

This positive effect of coherence, however, would not be expected with the situation model. While coherence helps the reader to build a strong textbase, it may hinder the formation of situation model connections (McNamara et al., 1996; McNamara, 2001). When texts are overly coherent, it may prevent high-knowledge readers from using their background knowledge to form situation model connections (McNamara et al., 1996). McNamara and her colleagues explained that incoherent text actually make high knowledge readers actively think about the organization of the text. Others have seen that recall of coherent text was related to reading comprehension, whereas recall of incoherent text was related to background knowledge (Voss & Silfies, 1996). Therefore, coherence actually decreases the need for background knowledge integration. The need for background knowledge for incoherent text is called "metacognitive gap sensing" (Britton et al., 1998). Only when the text has coherence gaps is background knowledge used to make inferences.

A third effect of coherence text may be the salience of the textbase information. Coherence, especially with the use of macrosignals, directs the reader to the important relationships in the text and, therefore, reduces the number of inferences a reader would have to make to form the textbase. For example, Lorch and Lorch (1996) examined the recall of information using a revised text that had half of the information signaled and compared it to the original text that had no signals. Information that was signaled in the revised text was remembered more frequently

than in the original text. However, the non-sigaled text in the revised text was recalled less than in the original text. Therefore, signals draw attention to sigaled information, while withdrawing attention from non-sigaled information.

When textbased links are sigaled in the text, they become more salient in the reader's mind. The rate that the reader pays attention to the signals should mainly affect the way the student rates the textbase links. Therefore, one main reader characteristic, the attention to signals, determines the rate that the reader makes these connections. However, when no signals are present in the text, multiple reader characteristics will determine the amount of textbase connections that are formed. First, the reader may need to rely upon background knowledge to form the connections and fill-in coherence gaps (Britton et al., 1998). Second, reading comprehension ability, especially inference generation, would need to be relied on to make these connections (Britton et al., 1998). Lastly, other minor cuing within the text may be relied on for information about the relationships among the concepts in the textbase (Meyer et al., 1980). Therefore, when macrosignals are not used, the reader's amount of background knowledge, reading comprehension ability, as well as attention to minor text cuing will all influence the textbase scores. Consequently, the items that form the textbase factor should form a more well-fitting model when macrosignals are used, indicating fewer reader characteristics driving the scores, than when no macrosignals are used.

In sum, coherence changes highlight the independence of the textbase and situation model connections in multiple ways. First, coherence positively affects textbase connections, while either not affecting or negatively affecting situation

model connections. Second, coherence, through the use of macrosignals, should increase the salience of the textbase connections. Coherence decreases the number of reader characteristics that influence how the textbase connections are formed. This may not be the case when connections are not signaled within the text. Therefore, the textbase factor should fit better when these connections are signaled within a text than when they are not.

Background Knowledge. The second step toward understanding the independence of these two text representation components was to document the effects of familiarity on the reader's ability to create a coherent textbase and situation model. Research has shown a positive effect of familiarity on situation model connections, as well as textbased connections (McNamara, 1994; McNamara et al., 1996).

The positive effect of familiarity on comprehension and text recall has been found in numerous studies (Britton & Gulgoz, 1991; Linderholm et al., 2000; Lorch & Lorch, 1996). Familiar text significantly increases recall more than unfamiliar text (Britton & Gulgoz, 1991). The familiarity of text has been varied by teaching readers the content before reading, (McKeown et al., 1992), assessing the amount of familiarity of subjects before reading (McNamara et al., 1996; Voss & Silfies, 1996), and by making explicit connections to background knowledge within a text (Vidal-Abarca & Sanjose, 1998). This positive effect of familiarity on recall has also been found with both textbase and situation model connections (McNamara et al., 1996).

This positive effect of familiarity on textbase connections, however, may not be found when text is too familiar. According to Kintsch (1998), "the mental text

representation is a mixture of text-derived and knowledge-derived information, not necessarily in equal parts. [In] extreme cases, the textbase or the situation model dominates to the exclusion of the other” (p. 104). Therefore, in an extreme case when text is very familiar, the situation model connections may dominate the knowledge network built during reading. After reading, readers are more likely to revert back to their strong background knowledge and overlook the connections described in the reading, especially when this knowledge is inconsistent with their prior knowledge (Woloshyn, Paivio, & Pressley, 1994). For example, in Appendix A, the words *purse* and *mermaid* may be very related in the text. In this text, a mermaid’s purse is the name of a shark egg. However, in a reader’s background knowledge he has never seen a picture or read about a mermaid carrying a purse. Therefore, he may not recall this relation after reading.

A third effect of familiar text may be the salience of the situation model information. Background knowledge facilitates the formation of situation model connections (Kintsch, 1998). This knowledge can be used in two ways. First, a reader can use background knowledge to fill-in coherence gaps in the text (Britton et al., 1998). Second, a reader can use background knowledge to make connections from coherent text to background knowledge, thus, integrating the two into a coherent whole. In both cases, the reader pays close attention to the information she already knew before reading the text. Background knowledge is always more integrated into memory than new knowledge (Kintsch, 1994). Also, low knowledge readers make few situation model connections. The more background knowledge one possesses, the more situation model connections can be created.

When text is familiar to the reader, one main reader characteristic, the amount of background knowledge activation, determines the rate that the reader makes these connections. This is evident in a well fitting situation model factor that is formed of situation model connections. However, when a text is unfamiliar, even more reader characteristics determine how the reader will form situation model connections. First, readers may rely upon their current background knowledge to answer questions about the unfamiliar topic. Second, when readers are very unfamiliar with a topic they may rely upon the structure of the words rather than their meanings. For example, some students may match similar sounding words in a rating task or match similar sounding question stems and item responses in a multiple-choice test. Finally, students guess on tests where they do not know the answer. Therefore, when unfamiliar text is presented, the reader characteristics such as reliance on faulty background knowledge, word structure, as well guessing all influence situation model connections. The greater number of reader characteristics determining how well a student forms the situation model decreases the fit of the situation model factor.

Similar to coherence changes, familiarity highlights the independence of the textbase and situation model connections. First, familiarity positively affects situation model connections, and may or may not affect textbase connections. Second, familiarity increases the salience of the situation model connections. Familiarity decreases the number of reader characteristics that influence how the situation model connections are formed. This may not be the case when these connections are found in unfamiliar text. Therefore, the situation model factor should fit better when these connections are from a familiar text than an unfamiliar text.

Conclusion

In conclusion, text coherence should increase textbase connections. However, coherence will not have as great an effect on situation model connections. Familiarity will have a positive effect on situation model connections, but will not have as great an effect on textbase connections. Text coherence will make the textbased links more salient and text familiarity will make the situation model links more salient. In turn, in a confirmatory factor analysis, the factor structure of the salient links will become clearer. These links should fall into a more one-dimensional factor structure when they are salient compared to when they are not salient.

Theoretical and Practical Significance

The theoretical purpose of this dissertation was to determine if individual components of knowledge constructed from text could be individually measured. Results from this dissertation should expand knowledge about the assessment of comprehension. However, this study has potential influences on instructional practice as well.

If the PCST was shown to measure knowledge structures students build from text, then teachers could use the PCST to assess students in their classrooms over time on their ability to form knowledge structures. Likewise, teachers could, over time, assess a student's ability to form specific components of this knowledge structure. With the results from the PCST the teacher may be able to tailor strategy instruction to rectify specific difficulties a student is having in building this knowledge structure. Finally, since the PCST can be applied to most texts, teachers

could modify this assessment to test student comprehension of individual classroom texts. Therefore, the PCST can be embedded within classroom instruction.

Unlike traditional comprehension measures that are paper-based, a correlation score from the PCST can be examined within minutes of the student taking the test. However, to give instantaneous individual component scores, more programming would be needed.

The use of the PCST as an instructional tool would require more research and development. However, the results from this dissertation may provide some recommendations for this instructional research. Therefore, the results of the current dissertation would not have an immediate impact on classroom instruction.

Hypotheses

Based on these theoretical assumptions about textbase and situation models, four research hypotheses were tested in this dissertation.

1. On the measure of textbase connections, students who read a text with macrosignals will outperform students who read a text without macrosignals.
2. On the measure of situation model connections, students who read a text on a familiar topic will outperform students who read a text on an unfamiliar topic.
3. Textbase scores will be more salient and will form a more one-dimensional factor structure for a text with macrosignals than a text without macrosignals.
4. Situation model scores will be more salient and will form a more one-

dimensional factor structure for a text on a familiar topic than a text on an unfamiliar topic.

Chapter 3: Method

Research Hypotheses

Four research hypotheses were tested in this dissertation.

1. On the measure of textbase connections, students who read a text with macrosignals will outperform students who read a text without macrosignals.
2. On the measure of situation model connections, students who read a text on a familiar topic will outperform students who read a text on an unfamiliar topic.
3. Textbase scores will be more salient and will form a more one-dimensional factor structure for a text with macrosignals than a text without macrosignals.
4. Situation model scores will be more salient and will form a more one-dimensional factor structure for a text on a familiar topic than a text on an unfamiliar topic.

Participants

Participating in this study were 236 ninth-grade students from 15 self-contained computer classes instructed by 2 computer teachers in a mid-Atlantic state school. The school can be described as an all male Catholic high school. Only students with permission from their parents participated in this study. These students were assessed with the National Educational Developmental Test during the year of the study. On average the students scored a national percentile of 62% on the social

studies reading subtest, 57% on the science reading subtest, 60% on the English usage subtest, and a 91% on the Mathematics usage subtest.

Materials

Six texts were used in this study: (a) a familiar text with macrosignals on the topic of the United States of America (5 pages) that contained macrosignals such as introductions, headings, bold words, and font sizes, (b) a familiar text without macrosignals on the topic of the United States of America (3 pages), (c) an unfamiliar text with macrosignals on the topic of Sri Lanka (5 pages), (d) an unfamiliar text without macrosignals on the topic of Sri Lanka (3 pages), (e) a control text on the topic of symbiosis (4 pages), and (f) a second control text on the topic of symbiosis (4 pages). Students assigned to the control text “e” were administered the same reading comprehension test assigned to students in the “a” and “b” groups and students assigned to control text “f” were assigned to the same reading comprehension test assigned to students in the “c” and “d” groups.

The America text and Sri Lanka text were selected because they were hierarchical in nature. Although other text structures could be selected, independent studies on the use of the PCST for measuring reading comprehension from text with these structures would need to be completed.

The two texts with macrosignals (a and b) were both five pages in length. Most of the text and pictures were pulled from two social studies trade books. The familiar text was extracted from the book *United States of America* written by Martin Hintz (2004) and the non-familiar text was extracted from the book *Sri Lanka* written by Krishnan Guruswamy (2002). The two experimental texts compiled for this

dissertation had one main title, two main section headings, and 14 sub-headings. Each text had seven black and white pictures with captions included on four of these pictures. Four introductory paragraphs were written for the macrosignals texts. The country and the text as a whole were introduced in the first paragraph on the first page. The next three introductory paragraphs introduced the main sections and were located at the beginning of each section, which was at the top of pages 2, 3, and 4 respectively. The familiar text with macrosignals had 987 words and 17 paragraphs. The unfamiliar text with macrosignals had 987 words and 15 paragraphs. Approximately 13 key words were italicized within each text.

The two texts with no macrosignals were both three pages in length. These texts were identical to the macrosignal texts; however, they did not include introductory paragraphs, headings, picture captions, or italicization of key words. These texts included the same seven pictures as the macrosignals texts. Without macrosignals the familiar text had 829 words and 12 paragraphs. The unfamiliar text without macrosignals had 827 words and 11 paragraphs.

The Passage Comprehension for Structured Text (PCST) assessment required the selection of key terms to symbolize the underlying conceptual structure of the text. Since the texts used in this investigation were hierarchical in nature, the words selected for the PCST represented this hierarchical structure. In a sense, the words were selected to form a shortened summary or outline of the text. These words were always selected directly from the text. In most cases, words were not selected if they were too difficult or meaningless to the reader.

The first word that was selected was the top of this hierarchy, which was the main topic of the reading. Nouns that were short but could summarize most of the information in which it represented were selected. For example, in the text about America, the word representing the main topic was “America.” The main word for the Sri Lanka text was “Sri Lanka.” The second two words were selected to represent the main sections within the text. Both texts were composed so that the two main sections in each reading would detail information about the history of the country and the modern everyday life of people in that country. In the America as well as the Sri Lanka texts the words “history” and “modern” were selected to represent the two main sections. These words were both representative of the sections and they were the actual words used in the headings as well as the text. The next words were selected to represent any subheadings under the main headings. The texts were written to have two main subheadings under the history section and three main subheadings under the modern section. Words were selected to represent these five subheadings. Next, an additional word was selected from one history section and two additional words were selected from the other history section. These words were selected to create the complexity of a four-layered hierarchical structure.

This structure was informed by the pilot study (see Appendix D) and changes were made to accommodate the limitations found. First, in the pilot study, two words were used to represent the section on modern everyday life. The words “modern” and “everyday” were both used. In the analysis of the pilot study, it was determined that students had more success with the word modern than the word everyday. Therefore the word everyday was dropped from the representation. Second, in the pilot study

the history section was represented by the word “early” in the America text and the word “ancient” in the Sri Lanka text. These words were used often in the original text. However, students had difficulty indicating relationships with these words. Therefore, both words were changed to the word history. Further, the text was changed to highlight the word history more than the previous words. Finally, in order to create a more coherent structure, more headings, picture captions, and preview statements were added.

In all, 11 words, representing the conceptual knowledge structure of the reading, were selected from the America text and 11 words were selected from the Sri Lanka text. Of the 11 words, one word symbolized the passage’s main theme, two words symbolized the main concepts in the reading, and eight words symbolized supporting phenomena to the two main concepts. In the America packet, *America* represented the main theme. The words representing the two concepts, found in the main headings of each section, were *history* (representing the past) and *modern* (representing the present). The words *settlers*, *Indians*, *independence*, *Stamp Act*, and *government* supported information to past America. The words *houses*, *food*, and *job* supported information to the present America. These words, in hierarchical order, can be seen in Figure 2. Figure 3 shows the hierarchical order of the words selected from the Sri Lanka text.

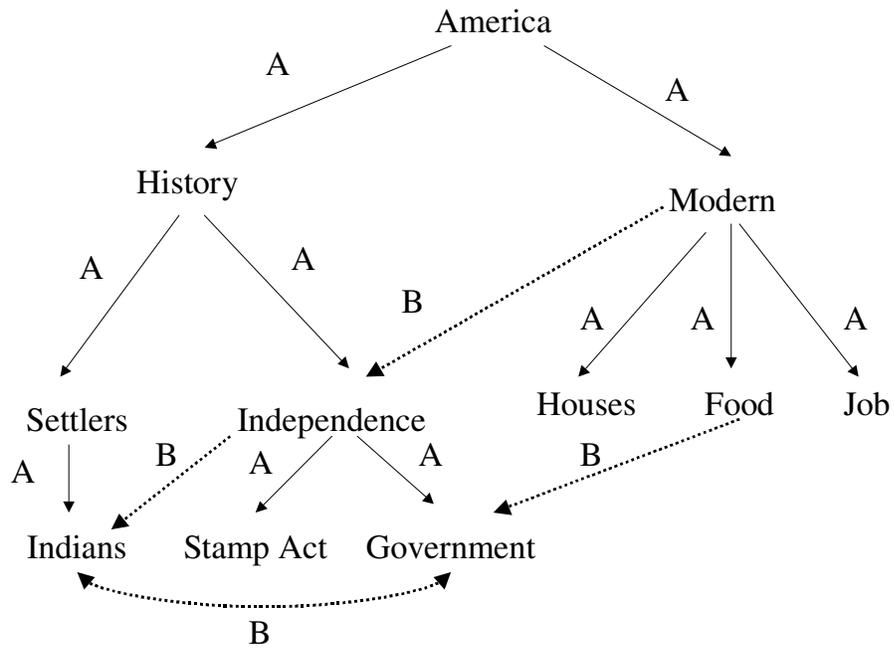


Figure 2. Representation of the text structure for America

A = textbase links, all links shown
 B = situation model links, three sample links shown

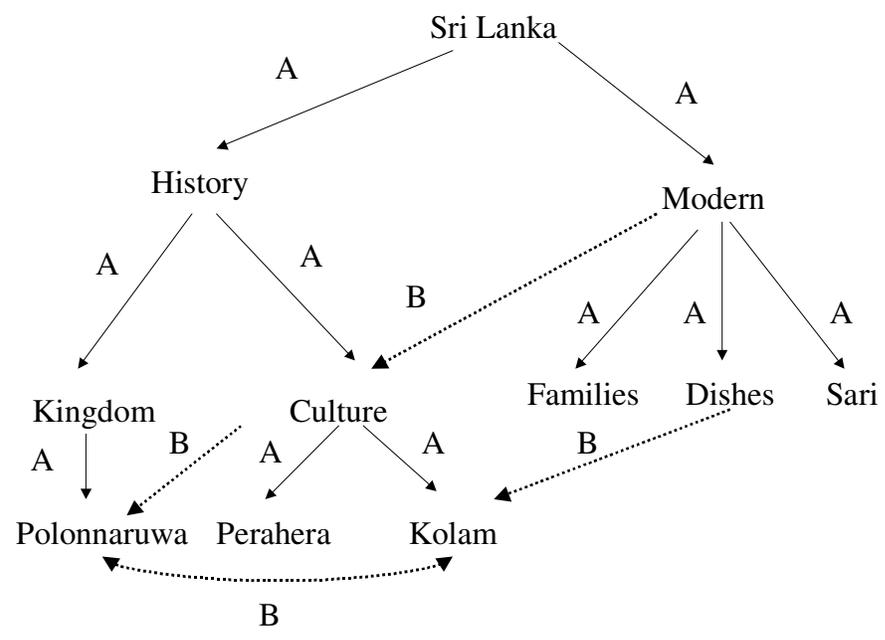


Figure 3. Representation of the text structure for Sri Lanka

Figures 2 and 3 are examples of an expert text representation, with a few example links. From the links, two comprehension components of the structure were identified as A (textbase) and B (situation model). Reading comprehension component A, representing the textbase, consisted of 10 links (e.g. Links labeled A). All of these are labeled on Figures 2 and 3. Component A was the sum of the links connecting the theme word with concept words (e.g., America/modern, expert rating of 9) and concept words with supporting phenomenon (e.g. modern/houses, expert rating of 9). Component B, represented the situation model, and included links involved in the distinction among concepts and phenomena (e.g., Links labeled B). These links were not explicitly stated in the text and, therefore, were informed through background knowledge. Twenty-eight links were identified. Four sample links are shown in Figures 2 and 3. Component B was the sum of the links connecting concept words with non-supporting phenomenon (e. g., modern/independence, expert rating of 1) and phenomenon with dissimilar phenomenon (e. g., food/government, expert rating of 1).

Measure: Passage Comprehension for Structured Text

To measure students' reading comprehension and the two components of textbase and situation model connections, students were administered the Passage Comprehension for Structured Text (PCST). This task took the entire class time (45 minutes) to complete.

This current study used a numerical similarity rating system similar to the method used by Britton and Gulgoz (1991). This method was selected since it

provided “rich knowledge structures specifying a relationship for each pair of concepts” (Ferstl & Kintsch, 1999). Second, the similarity ratings were easy to collect and to use in order to assess knowledge structures. Third, numeric ratings were used in Britton and Gulgoz (1991), one of the few research papers to use proximity data with nonfiction text. Therefore, since numeric ratings were used successfully before, I believed that it would work in a similar situation. Finally, since a relationship or lack of relationship is specified for every pair of concepts, these relationships could be compared among all the students.

The numeric rating system and the pathfinder procedure used to collect proximity data has not been previously studied with students in high school. This is a new attempt to generalize this procedure to this population.

Definition

Passage Comprehension referred to a student’s ability to comprehend a short passage and rate the relatedness of key words from that passage.

Procedure

Students were randomly assigned to one of six texts. These texts were located on their assigned desks along with a student assent form and a rating sheet. Once the students were at their desks, the teacher read the information located on the student assent sheet. The students were asked to sign their student assent sheet. Afterwards, the students were asked to provide information at the top of the student rate sheet, which included their name, date, classroom number, and computer number. Once all the students filled-in their information, the teacher directed the students to turn over

their assigned reading. At this point the teacher said, “You will have 15 minutes to read this text. As you read, think about the big ideas, major concepts, and important relationships among the information in the reading. Think about how these big ideas relate to each other.”

Students were instructed not to mark on their materials. Students who finished early were directed to look through their packets again. Once all students finished reading their passages, the teacher directed the E and F students to write important key words from their reading onto a blank paper provided at their computer. The A through D students were directed to place their readings face down on the computer stand and to look at the practice sheet which was copied on the opposite side of the rating sheet. The teacher said:

Now we will see how much of the reading you can remember. But first, we will do three practice items. Turn your rating sheet over to see the practice sheet. It should look like this (The teacher held up the practice sheet and made sure all students were looking at the correct side).

The teacher read through the practice items as follows:

What words do you see at the top of the page? How do you think these words are related? Are Wolf and Fangs related (wait for student response)? Yes, A wolf has fangs. You would give those words that are related a 9. Circle the 9 on your paper. Are Butterfly and Fangs related (wait for student response)? No, A butterfly does not have fangs. You would give those words that are not related a 1. Circle the 1 on your paper. Are Wolf and Butterfly related (wait for student response)? Yes, A wolf and a butterfly are both animals, however

they are also a bit different. You would give those words that are a little bit related a 5. Circle the 5 on your paper.

The teacher then directed the A and B students to the computers, while an assistant directed the C and D students. For every other class, the teacher and assistant reversed their assignments. Once the students were looking at their screens, the teacher or assistant made sure that the students had the correct screen on their monitors. The teacher then said:

Now we will do the same thing with words from your reading. As you rate the words, think about what you just read ---the big ideas, the major concepts, and relationships from that reading. Click the space bar one time. You should see 11 words from your reading. Please point to the words as I read them. (For America: Job, Settlers, Houses, Indians, Stamp Act, History, Food, America, Government, and Independence). Look at these words all together. Think about how these words are related. Now tap the space bar one more time. You will see two words. Decide if these words are very related, moderately related, or not related, by typing a 1, 5, or 9. When you are sure of your answer, type in the number, and hit the space bar for your next set of words. Keep rating the words until the computer tells you to stop. Remember to think about the big ideas explained in your reading.

Once the A, B, C, and D students had started their task, the teacher and the assistant collected the E and F key-word papers, asked the students to place their readings face down on their computer stands, and read their directions. The directions were similar to that of the A through D students except the E and F students were

directed to think about important relationships of the words rather than directing them to think about what they just read.

As students finished their tasks, the assistant recorded each student's file number on the rating sheet. When all the students had finished rating their words, they turned in their answer sheets and student assent forms and received a candy bar.

Validity and Reliability

Internal constancy alphas were established for the two forms (familiar and non-familiar) of the PCST in the pilot study located in Appendix D. The internal consistency reliabilities for the scores on the textbase subscale were .87 for *America* and .63 for *Sri Lanka*. The internal consistency reliabilities for the scores on the situation model subscale were .87 for *America* and .62 for *Sri Lanka*. The internal consistency reliabilities for the updated PCST forms can be examined in Chapter 4.

Previous construct validity for proximity data gathered with numerical ratings was found in a study with third-grade students. Guthrie et al. (2004) administered a passage comprehension measure similar to the PCST assessment. Guthrie et al. found high correlations between the correlation score on the passage comprehension assessment with an essay reading comprehension measure (.82), as well as with the Gates-MacGinitie Reading Comprehension subtest (.82).

Four experts established the validity of the expert model. The four experts were two social studies teachers and two history graduate students, who were asked to perform three tasks. All three tasks were performed with both the *America* text and the *Sri Lanka* texts and were counterbalanced for order among the experts. First, each expert was asked to rate the expert model on a 1 -10 scale, from not at all like the

representation of the text to exactly like the representation of the text. Findings are found in Table 2. The mean score for the expert model was 7 for the *America* text and 7.5 for the *Sri Lanka* text. Next, the experts were asked to form their own representation out of the words used in the experimenter representation. Of the 10 links found in the expert model, the *America* text had a mean of 7.75 number of common links and the *Sri Lanka* text had a mean of 9.25 number of common links. Afterwards, they were asked to form their own expert representations. The number of concepts selected as key terms from the original eleven words was 10.00 for the *America* text and 9.75 for the *Sri Lanka* text. These results indicate that the expert structures do a fairly reliable job describing the structure of the text. Also, the Sri Lanka text expert structure seemed as good as the American expert structure.

Table 2

Expert Ratings for Each Text

	Teacher 1	Teacher 2	Student 1	Student 2	Total
<i>America</i>					
Rating (1-10)	4	8	9	7	7.00
Links (10)	6	8	8	9	7.75
Concepts (11)	10	9	11	10	10.00
<i>Sri Lanka</i>					
Rating (1-10)	7	8	10	5	7.50
Links (10)	9	10	9	9	9.25
Concepts (11)	9	9	11	10	9.75

Scoring

Seven scores were obtained from the rating data. The first score was the raw rating from each student for each pair of words. Students were directed to use the ratings of 1 for not related, a 5 for a little bit related, and 9 for very related. Although students were directed to use only the ratings 1, 5, and 9, a few students used other numbers on the continuum. In order to maintain some order, without missing data, raw scores were recoded such that: (a) 1, 2, and 3 were coded as a 1, (b) 4, 5, and 6 were coded as a 5, and (c) 7, 8, and 9 were coded as a 9.

The second score was the difference score and was computed by taking the difference of the expert and the student rating on each pair of items. Since the

students only use the 1, 5, and 9 scores, the difference scores included -8 , -4 , 0 , 4 , and 8 . Therefore, the range of scores was from -8 to $+8$. A score of zero indicated no difference between the student and the expert. To the extent a score deviates from zero indicated the difference of the student from the expert. Although many scores could be used to determine difference from an expert, such as Euclidian distance, I used the difference between the expert and the student ratings since it represents the simplest score, and can permit a reader to connect raw data to the scale of measurement. The third score was the distance score and was the absolute value of the difference score. These scores included 0 , 4 , and 8 . This score indicated the distance a student is from the expert without indicating the direction. The fourth score was the reversal of the distance score. The score was reversed so that high scores represented high similarity to the expert and low scores represented low similarity to the expert. These scores still included a 0 , 4 , and 8 . The fifth score was the textbase score and was computed by summing the 10 textbase reversed distance scores. A list of the textbase connections can be found in Appendix B. These scores ranged from 0 to 80 (10 items by score of 8). A score of 0 indicated complete difference from the expert and a score of 80 indicated complete similarity to the expert. In the pilot study, 11 scores were used to represent the textbase connections. The results of the pilot study, found in Appendix D, indicated that students ranged on this measure from 0 to 72 , with a mean of 39.45 and a standard deviation of 20.43 . The sixth score was the situation model score and was computed by summing the 28 situation model reversed distance scores. A list of the situation model connections can be found in Appendix C. These scores ranged from 0 to 224 (28 items by score of 8). A score of 0 indicated

complete difference from the expert and a score of 224 indicated complete similarity to the expert. In the pilot study, 30 scores were used to represent the situation model connections. The results of the pilot study, found in Appendix D, indicated that students ranged on this measure from 92 to 240, with a mean of 155.47 and a standard deviation of 36.35. The final score was the correlation score. This score was the correlation among the 55 total ratings of all the possible word pairs by the student with the 55 ratings of the expert. The 17 scores not labeled as textbased or situation model were rated as “a little bit similar” and given an expert model score of 5. The correlation score represents the student’s general similarity with the expert. This score was not analyzed in the pilot.

Administrative Procedures

The total passage comprehension assessment was administered in one class period (45 minutes). All students in each classroom of two teachers during a 2-day duration were assessed. Each teacher administered the reading section of the PCST as well as the computer section of the PCST. Two research assistants, a former middle school teacher and the researcher, were assigned to assist one of the two teachers. At three separate times (2 in Day 1 and 1 in Day 2) both teachers were administering the test at the same time in two labs. The sequence of events was:

1. Hand out Student Permission forms and rating sheets (5 minutes).
2. Students read passage (15 minutes).
3. Teacher reads through practice items while students who read E and F write key words (5 minutes).

4. Teacher reads through practice items with E and F students while A, B, C, and D students start their ratings (5 minutes).
5. Students complete their tasks (15 minutes).

The researcher met with the teachers before the data collection. During this meeting, the experimenter reviewed the directions. Also, the experimenter familiarized the teachers with the computer program. The teachers were told that during the administration of this study no student would be allowed to start any other activity during the study (reading a book or completing homework). The teacher was not allowed to help the students read the test. During the computer administration, the teacher was allowed to read the words from the screen, but was not allowed to give any other information about the words.

Chapter 4: Results

Preliminary Analyses

The dependent variables examined in this study were textbase and situation model representation scores as well as correlation scores. The independent variables examined in this study were the topic familiarity (1 = familiar, 0 = not familiar), and the text type (2 = macrosignals, 1 = no macrosignals, 0 = control text).

Descriptive Statistics

First I examined the number of students in each group. Due to class absences, some readings were assigned to more students. Table 3 shows the distribution of students in these groups. To even the groups, students were randomly dropped from conditions B (4), C (10), and D (2) to make the numbers similar to condition A. Eight students were also dropped from condition E to make it similar to condition F. The random dropping of students from these conditions was implemented through SPSS.

Table 3

The number of student in each condition before equating the groups

Reading	Number of Students
A	42
B	46
C	52
D	44
E	41
F	33

Next, I examined the data to identify outliers. Multivariate Analysis of Variance (MANOVA), used in the next stage of the analysis, is sensitive to outliers since they may impact the Type I error of the MANOVA. Therefore, it was important to identify and remove any extreme outliers in the data. This identification was processed in four steps. First, I examined the outliers within the groups sorted by text type using a box-plot formed in SPSS. Textbased scores of the students sorted by text type indicated four outliers, two students with a reading C, one with a reading F, and one with a reading E. Situation model scores of the students sorted by text type indicated three outliers, two students with a reading A, and one student with a reading D. Second, I examined the outliers within groups sorted by familiarity. Textbased scores of the students sorted by familiarity indicated one outlier with reading D. Situation model scores of students sorted by familiarity indicated two students whom

were already indicated within the text type outlier analysis. In all, eight students were identified and deleted from the sample. In order to balance the groups, two reading B students were randomly deleted as well.

The means, standard deviations, and numbers of students in the six text conditions are presented on Table 4. The correlation matrix of these variables is shown in Table 5. As can be seen from Table 4, students in the macrosignals groups had consistently higher means on the textbase and situation model scores, as well as the correlation score. Students in the familiar text groups had consistently higher means on the textbase and situation model scores, as well as the correlation score. From Table 5, one can notice that familiarity had a higher correlation with situation model scores than textbase scores. Also, macrosignals had a higher correlation with textbase scores than situation model scores. There seemed to be a slightly negative correlation between the textbase scores and the situation model scores. In the pilot, found in Appendix D, this negative correlation was more pronounced.

Table 4

*Means and Standard Deviations of Textbase and Situation Model**Connections and Correlation Scores for Each Text*

Text	Textbase	Situation Model	Correlation
Familiar Macrosignals (N = 40)			
Mean	62.80	150.30	.51
SD	10.88	28.22	.13
Familiar No Macrosignals (N = 40)			
Mean	58.70	148.50	.43
SD	12.63	26.89	.20
Familiar Control (N = 33)			
Mean	61.70	134.06	.42
SD	9.49	28.98	.12
Unfamiliar Macrosignals (N = 40)			
Mean	56.20	122.90	.27
SD	10.56	26.17	.12
Unfamiliar No Macrosignals (N = 40)			
Mean	51.20	119.40	.19
SD	12.43	27.03	.13
Unfamiliar Control (N = 33)			
Mean	41.82	122.42	.06
SD	11.00	28.86	.16

Table 5

*Intercorrelations Among Familiarity, Macrosignals, Situation**Model Connections and Textbase Connections in the Treatment Groups*

	1	2	3	4
1. Familiarity	-			
2. Macrosignals	.00	-		
3. Situation Model	.47**	.04	-	
4. Textbase	.29**	.19*	-.07	-

Note. * $p < .05$, ** $p < .01$

Reliability Analysis

Internal consistency alphas, performed on the reversed distance scores, were computed for the two forms (America and Sri Lanka) of the passage comprehension assessment. Distance scores were the absolute difference between the student's score and the expert score for each word-pair. The internal consistency reliabilities for the scores on the two subscales (Textbase and Situation model) of the two forms of the PCST (America and Sri Lanka) were calculated. Reliability coefficients for scores on the 10 textbase links were .61 for America and .48 for Sri Lanka. Reliability coefficients for scores on the 28 situation model scores were .77 for America and .69 for Sri Lanka. In the following results, all 10 textbased reversed distance scores were used to calculate the textbased total score and all 28 situation model reversed distance scores were used to calculate the situation model total score. The 10 textbased reversed distance scores and the 28 situation model reversed distance scores were also used in the Confirmatory Factor Analysis (CFA).

Multivariate Analysis of Variance

To examine the first two hypotheses of this dissertation a MANOVA was conducted with text type and familiarity as independent variables and textbased connections and situation model connections as dependent variables. The MANOVA examined both the main effects of these independent variables on each dependent variable, as well as interaction effects. Compared to univariate ANOVAs or t-tests, the MANOVA was able to examine the combined effects of the dependent variables as well as control for experiment wide error rate.

Research Hypothesis 1

The first hypotheses of this dissertation stated that on the measure of textbase connections, students who read a text with macrosignals would outperform students who read a text without macrosignals. The main effect of text type on textbased scores was examined in order to test the first hypothesis. This analysis indicated that the students who received macrosignals ($M = 59.50$, $SD = 11.16$) outperformed students without macrosignals ($M = 54.95$, $SD = 13.01$), $F(1,156) = 6.09$, $p < .05$, with an effect size of .38. The main effect of text type on situation model scores was examined next. However, there was no significance between students who received macrosignals ($M = 136.60$, $SD = 30.35$) and students who did not receive macrosignals ($M = 133.95$, $SD = 30.53$), $F(1,156) = .38$, $p > .05$, with an effect size of .09.

Research Hypothesis 2

The second hypothesis stated that on the measure of situation model connections, students who read a text on a familiar topic would outperform students who read a text on an unfamiliar topic. The main effect of familiarity on situation model scores was examined in order to test the second hypothesis. This analysis indicated that the students who received the familiar text ($M = 149.40$, $SD = 27.40$) outperformed students who received the unfamiliar text ($M = 121.15$, $SD = 26.50$), $F(1,156) = 43.50$, $p < .001$, with an effect size of 1.05. The main effect of familiarity on textbased scores was examined next. This analysis indicated that the students who received the familiar text ($M = 60.75$, $SD = 11.89$) outperformed students who

received the unfamiliar text ($M = 53.70$, $SD = 11.73$), $F(1,156) = 14.62$, $p < .001$, with an effect size of .60.

Finally, I examined interaction effects between macrosignals and familiarity. The interaction was not significant for either textbase scores, $F(1,156) = .06$, $p = .81$ or situation model scores, $F(1,156) = .04$, $p = .84$.

General Structure Analysis

To examine whether students had a similar general structure to the expert, correlation scores were analyzed. Students in the macrosignals condition ($M = .39$, $SD = .17$) had statistically higher similarity to the expert than the non-signals condition ($M = .31$, $SD = .21$), $t(158) = 2.6$, $p < .01$, with an effect size of .41. Students in the familiarity text condition ($M = .47$, $SD = .17$) had statistically higher similarity to the general expert model than the students in the unfamiliar text condition ($M = .23$, $SD = .13$), $t(158) = 10.01$, $p < .001$, with an effect size of 1.52.

Control Group Analysis

To examine whether students were able to do well on the assessment without having read the text, control groups were examined. Students in the control condition were compared to students in the macrosignals and no macrosignals conditions. In all three conditions, familiarity was collapsed so that half of the students in each condition received the familiar text and half of the students received the unfamiliar text. An analyses of variance revealed a main effect of text type on textbase scores, $F(2,223) = 6.80$, $p = .00$, $\eta_p^2 = .06$, and a main effect of text type on correlation, $F(2,223) = 9.50$, $p = .00$, $\eta_p^2 = .08$ but text type did not have a main effect on situation

model scores, $F(2,223) = 1.43$, $p > .05$, $\eta_p^2 = .01$. The Scheffé post hoc multiple comparisons test was used to examine the differences between the conditions for the textbase and correlation scores. The results of the Scheffé-test on textbase knowledge scores revealed a statistically significant difference between the control text ($M = 51.76$, $SD = 14.29$) and the macrosignals text ($M = 59.50$, $SD = 11.16$), ($p < .01$), with an effect size of .60. The difference between the control text and the nomacrosignal text ($M = 54.95$, $SD = 13.01$), was not statistically significant, ($p > .05$), with an effect size of .23. The results of the Scheffé-test on the correlation scores revealed that there was a statistically significant difference between the control text ($M = .24$, $SD = .23$) and the macrosignals text ($M = .39$, $SD = .21$), ($p < .001$), with an effect size of .72. The difference between the control text and the nomacrosignal text ($M = .31$, $SD = .21$), was not statistically significant, ($p > .05$), with an effect size of .31.

Confirmatory Factor Analyses

Research Hypothesis 3

The third hypotheses of this dissertation was that textbase scores will be more salient and will form a more one-dimensional factor structure when students read text with macrosignals than without macrosignals. To test this hypothesis, confirmatory factor analysis (CFA) was used to determine whether the fit of the 1 factor model made of the 10 textbased scores was different for students who read macrosignal passages compared to students who read passages without macrosignals. That is, whether there was any additional model fit with two groups selected compared to

having no groups selected. To determine if there was a difference between groups, it was first determined whether adding the additional group information, the dummy coded variable for groups, would improved the fit of the factor compared to having no group information.

Three models were compared to determine the difference between the two groups. A CFA was first performed on Model 1, the homogeneous model. In the first model, all the data was combined without distinguishing groups. A second CFA was performed on Model 2, the partial heterogeneous model. In the second model, group indicator residual variances and covariances were allowed to vary, but the factor loadings were held constant between groups (Muthén & Muthén, 2005). Finally, a CFA was performed on Model 3, the fully heterogeneous model. In the third model, group indicator residual variances and covariances as well as factor loadings were allowed to vary across groups. The Akaike Information Criterion (AIC) was examined to determine the best fitting model. If the third model had the lowest AIC it would indicate that there was a significant difference between the factor loadings of the two groups.

The results of these three models are summarized in Table 6 in the column labeled one factor, Macrosignals. Model 1, representing the homogeneous data set, had an AIC of 7614.07, which was the worst fitting model. Model 2, the partially heterogeneous model, had an AIC of 7603.39, which was the best fitting model. Model 3, the fully heterogeneous model, had a slightly higher AIC than Model 2, indicating a slightly worse fit. The results of these models indicate a better fitting

model when group membership is taken into account; however, factor loadings are held constant.

The fit of the 1-factor model was examined individually for the 80 students in the macrosignals condition and the 80 students in the no macrosignals condition. The Comparative Fit Index (CFI) of the individual CFAs was used to determine strength of fit. The results of these analyses can be seen in Table 7. As can be seen from this table, Macrosignals had a higher CFI (1.0) than no macrosignals (.72). In addition, the macrosignals condition had a non-significant chi-square, indicating that the model does not show disparity from the observed data. However, the chi-square of the no macrosignals condition was significant.

The factor loadings for these CFAs can be found in Table 8 under the headings “Loadings Macrosignals” and “Loadings No Macrosignals.” As can be seen from these loadings, the macrosignals condition had three loadings that reach significance. In the no macrosignals condition, no factor loadings reached significance.

In summary, the results indicate that the two groups were not statistically different in their factor loadings. However, the residual variance between the groups was shown to vary. However, it is not clear from the data whether the residual variance was due to measurement error or to other unmeasured factors. The individual analysis of the conditions provided some, although limited evidence, suggesting that the 1-factor textbase model of the macrosignals condition had a better fit than the 1-factor textbase model of the no macrosignals condition.

Research Hypothesis 4

The fourth hypothesis of this study was that situation model scores will be more salient and will form a more one-dimensional factor structure when students are tested on a familiar topic than an unfamiliar topic. This hypothesis was tested in the same manner as the previous hypothesis. In this analysis, I used confirmatory factor analysis (CFA) to determine whether the fit of the 1 factor model made of the 28 situation model scores was different for students who read familiar text compared to students who read unfamiliar text. To determine if there was a difference between groups, I had to first determine whether adding the additional group information, the dummy coded variable for groups, would improved the fit of the factor compared to having no group information.

The three models compared to determine the difference between the familiar text and unfamiliar text conditions were similar to the models compared in the analysis of Hypothesis 3. These included the homogeneous model, the partially heterogeneous model, and the fully heterogeneous model. If the fully heterogeneous model is shown as the best fitting model, as indicated by having the lowest AIC, then it could be concluded that the conditions were statistically significantly different in their factor loadings.

The results of these three models are summarized in Table 6 in the column labeled one factor, familiarity. Model 1, representing the homogeneous data set, had an AIC of 21,889.23, which was the worst fitting model. Model 3, the fully heterogeneous model, had an AIC of 21,367.15, which was the best fitting model.

The results of these models indicated a better fitting model when group factor loadings were allowed to vary.

The fit of the 1-factor situation model was examined individually for the 80 students in the familiar text condition and the 80 students in the unfamiliar text condition. The Comparative Fit Index (CFI) of the individual CFAs was used to determine strength of fit. The results of these analyses can be seen in Table 7. As can be seen from this table, familiar text had a higher CFI (.54) than unfamiliar text (.40). The results of these analyses show that these two groups were different and that the 1-factor model fit the familiar text group better than the unfamiliar text group.

The factor loadings for these CFAs can be found in Table 8 under the headings “Loadings Familiar” and “Loadings Unfamiliar.” These loading show that the familiar condition had 17 loadings that reach significance. In the unfamiliar condition, 15 factor loadings reached significance.

Further Analysis: The Two Factor Model

Additional analyses were completed to compare the strength of the two-factor structure, textbase and situation model, with the macrosignals and no macrosignals conditions as well as the familiar and unfamiliar text conditions. Although these analyses were not particularly crucial for testing hypotheses three and four, I believed that they provide interesting findings to guide future research in this area. Figure 4 depicts the two-factor model that was tested.

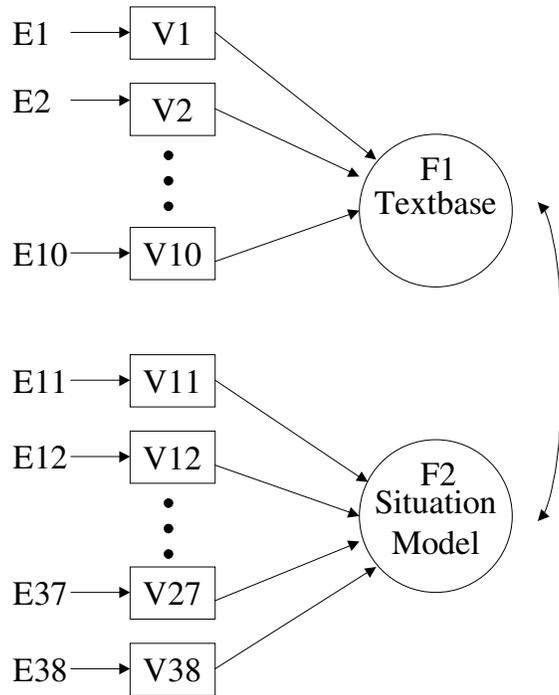


Figure 4. Representation of the two-factor model

Macrosignals effect on the two-factor model. I used confirmatory factor analysis (CFA) to determine whether the fit of the 2 factor model, made of the 10 textbased scores and the 28 situation model scores, was different for students who read macrosignal passages compared to students who read passages without macrosignals. To determine if there is a difference between groups, I had to first determine whether adding the additional group information, the dummy coded variable for groups, would improved the fit of the 2 factors compared to having no group information.

Three models were compared to determine the difference between the two groups on the fit of the two-factor model. A CFA was performed on the first model,

the homogeneous model. In the first model, all the data was combined without distinguishing groups. A second CFA was performed on the second model, the partial heterogeneous model. In the second model, group indicator residual variances and covariances were allowed to vary, but the factor loadings were held constant between groups. Finally, a CFA was performed on the third model, the fully heterogeneous model. In the third model, group indicator residual variances and covariances as well as factor loadings were allowed to vary across groups. The AIC was examined to determine the best fitting model. If the third model had the lowest AIC it would indicate that there was a significant difference between the factor loadings of the two groups.

The results of these three models are summarized in Table 6 in the column labeled two factors, macrosignals. Model 3, representing the homogeneous data set, had an AIC of 29,533.29, which was the worst fitting model. Model 2 had an AIC of 29501.97, which was the best fitting model. Model 3 had a slightly higher AIC than Model 2, indicating a slightly worse fit. The results of these models indicate a better fitting model when group membership is taken into account; however, factor loadings are held constant.

The fit of the 2-factor model was examined individually for the 80 students in the macrosignals condition and the 80 students in the no macrosignals condition. The Comparative Fit Index (CFI) of the individual CFAs was used to determine strength of fit. The results of these analyses can be seen in Table 7. As can be seen from this table, Macrosignals had a higher CFI (.42) than no macrosignals (.40).

Familiarity effect on the two-factor model. I used confirmatory factor analysis (CFA) to determine whether the fit of the 2-factor model, made of the 10 textbased scores and the 28 situation model scores, was different for students who read familiar text compared to students who read unfamiliar text. To determine if there was a difference between groups, I had to first determine whether adding the additional group information, the dummy coded variable for groups, would improved the fit of the 2 factors compared to having no group information.

The three models compared to determine the difference between the familiar text and unfamiliar text conditions, were similar to the models compared in the previous analyses. These included the homogeneous model, the partially heterogeneous model, and the fully heterogeneous model. If the fully heterogeneous model is shown as the best fitting model, as indicated by having the lowest AIC, then it could be concluded that the conditions were statistically significantly different in their factor loadings.

The results of these three models are summarized in Table 6 in the column labeled two factors, Familiarity. Model 1, representing the one group data set, had an AIC of 29,504.85, which was the worst fitting model. Model 1 had an AIC of 28,909.02, which was the best fitting model. The results of these models indicate a better fitting model when group factor loadings were allowed to vary.

The fit of the 2-factor model was examined individually for the 80 students in the familiar text condition and the 80 students in the unfamiliar text condition. The Comparative Fit Index (CFI) of the individual CFAs was used to determine strength of fit. The results of these analyses can be seen in Table 7. As can be seen from this

table, familiar text had a higher CFI (.44) than no macrosignals (.27). Results show that these two groups are different and that the 2-factor model fit the familiar text group better than the unfamiliar text group.

Table 6

Model fit for Models 1, 2, and 3 in the Textbase and Situation Model Analyses

Model	One Factor		Two Factors	
	Macrosignals	Familiarity	Macrosignals	Familiarity
1: Homogeneous				
Chi-square	42.24	706.32*	1180.58*	1180.58*
df	35	350	664	664
AIC	7614.07	21889.23	29504.85	29504.85
2: Partially				
Heterogeneous				
Chi-square	93.40	1195.94*	2193.10*	2198.134*
df	79	727	1364	1364
AIC	7603.39	21418.59	29501.97	28938.93
3: Fully				
Heterogeneous				
Chi-square	81.37	1090.50*	2152.42*	2096.22*
df	70	700	1328	1328
AIC	7609.36	21367.15	29533.29	28909.02

Note. AIC = Akaike Information Criterion; Bold text represents the best fitting model according to the AIC.

*p < .01

Table 7

Model fit for separate models

Model	Chi-square	CFI	RMSEA
Textbased Model			
Macrosignals	31.37	1.0	.00
No Macrosignals	50.00*	.72	.07
Situation Model			
Familiar	529.90**	.54	.08
Unfamiliar	560.60**	.40	.09
Two Factor Model			
Macrosignals	1082.73**	.42	.09
No Macrosignals	1069.69**	.40	.09
Familiar	992.75**	.44	.08
Unfamiliar	1103.47**	.27	.09

Note. CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation

* $p < .05$, ** $p < .01$

Table 8

Factor Loadings for Separate Models

<u>One Factor Textbase</u>			<u>One Factor Situation Model</u>		
Item	Loading Macrosignals	Loading No Macrosignals	Item	Loading Familiar	Loading Unfamiliar
1	.37	.10	1	.36	.40
2	.25	.48	2	.61*	-.16
3	.17	.13	3	.40*	.37*
4	.12	.20	4	.39*	.02
5	.10	.46	5	.25	.15
6	-.05	.42	6	.62*	.35*
7	.32	.42	7	.47*	.22
8	.54*	.46	8	.54*	.36*
9	.62*	.50	9	.37*	.43*
10	.44*	.44	10	.33*	.40*
			11	.09	.52*
			12	.62*	-.18
			13	.55*	-.13
			14	.51*	.15
			15	.22	.62*
			16	.38*	.51*
			17	.19	.41*
			18	.60*	.55*
			19	.12	.33*
			20	.51*	.08
			21	.30*	.27
			22	.34*	.31*
			23	.41*	.38*
			24	.10	-.02
			25	-.14	.57*
			26	.13	.23
			27	.06	.29*
			28	.08	-.08

* p < .05

Research Hypotheses 3 and 4 Assuming Ordinal Data

Recently the psychometric community has made recommendations on the appropriateness of using non-ordinal tests if the data may be assumed to be ordinal. In the Pathfinder literature, a few researchers have assumed the ratings to be ordinal in

nature (Goldsmith & Johnson, 1990). Therefore, in an effort to assure that treating data as scaled data would not produce different effects and to further examine the nature of the PCST, the previous analyses on Hypotheses 3 and 4 were re-analyzed by treating the ratings as ordinal data.

The findings of these analyses can be found on Tables 9 and 10. As can be seen on Table 9, the output from an analysis with ordinal data does not produce an AIC. Further, the comparison of these chi-squares cannot be directly tested with the chi-square and degrees of freedom (df) since “the conventional approach of taking the difference between the chi-square values and the difference in the degrees of freedom is not appropriate because the chi-square difference is not distributed as chi-square” (Muthén & Muthén, 2005, p. 378). Therefore, for this analysis, I used a corrected chi-square through the two-step DIFFTEST option in Mplus. In multiple-group analysis using Mplus a user could compare two models with similar baseline models, which included models 2 and 3. During the analysis, one item from the textbase model and four items from the situation model had to be dropped due to convergence problems. The difference between model 2 and model 3 of the familiar text and the unfamiliar text conditions was statistically different for both the one, $\chi^2(14) = 62.09, p < .01$ and two factor analyses, $\chi^2(20) = 66.32, p < .01$. Also, the difference between models 2 and model 3 of the macrosignal condition and the no macrosignal condition was not significantly different for both the one and two factor analyses.

The fit of the 1-factor and 2-factor models were examined individually for the 80 students in the macrosignals condition, the 80 students in the no macrosignals condition, the 80 students in the familiar condition, and the 80 students in the

unfamiliar condition. The results of these analyses can be seen in Table 10. Using the CFI to determine strength of fit, the 1-factor textbase model fit better with the macrosignals condition (1.0) than the no macrosignals condition (.86). The 1-factor situation model fit better with the familiar text condition (.78) than the unfamiliar text condition (.55). The 2-factor model fit best with macrosignals condition (.75) compared to the no macrosignals condition (.72). Finally, the 2-factor model fit best with the familiar text condition (.72) compared to the unfamiliar text condition (.46).

The findings from these analyses are similar to the findings from the previous analyses in multiple ways. First, when examining the difference between the models, the fully heterogeneous model was a statistically better fit than the partially heterogeneous model for the familiar and unfamiliar text conditions. Second, the two heterogeneous models were not significantly different, which was similar to the finding with the macrosignals and no macrosignals condition with the non-ordinal data. Third, while examining the fit of the individual group analysis, the macrosignals condition fit better than the no macrosignals condition; however, this finding was more pronounced in with the 1-factor model. Finally, while examining the fit of the individual group analysis, the familiar text condition fit better than the unfamiliar text condition for both the 1 and 2-factor models.

Table 9

Model fit for Models 1, 2, and 3 in the Textbase and Situation Model analyses with ordinal data

Model	One Factor		Two Factors	
	Macrosignals	Familiarity	Macrosignals	Familiarity
1: Homogeneous				
Chi-square	23.25	148.80	166.59	166.59
df	20	70	92	92
CFI	.95	.79	.76	.76
RMSEA	.03	.08	.07	.07
2: Partial Heterogeneous				
Chi-square	45.45	189.33	163.22	195.28
df	41	90	108	108
CFI	.93	.49	.75	.47
RMSEA	.04	.12	.08	.10
3: Fully Heterogeneous				
Chi-square	102.60	165.66	163.82	182.61
df	42	89	107	107
CFI	.96	.61	.74	.54
RMSEA	.03	.10	.08	.09

Note. CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation.

Table 10

Model fit for separate models with ordinal data

Model	Chi-square	CFI	RMSEA
Textbased Model			
Macrosignals	16.70	1.0	.00
No Macrosignals	23.95	.86	.07
Situation Model			
Familiar	67.16*	.78	.08
Unfamiliar	82.03**	.55	.10
Two Factor Model			
Macrosignals	80.73*	.75	.08
No Macrosignals	83.19**	.72	.09
Familiar	77.56*	.72	.08
Unfamiliar	95.16**	.46	.09

Note. CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of

Approximation

* $p < .05$, ** $p < .01$

Chapter 5: Discussion

Two major goals directed this dissertation. The first goal of this dissertation was to determine whether theoretical components of reading comprehension were measured by the Passage Comprehension for Structured Text (PCST) assessment. To examine these components, the sensitivity of each component to text manipulations was measured. To this end, a MANOVA was conducted to measure the main effects of coherence and familiarity on the components of textbase and situation model connections. If the measure were valid, I expected it to be sensitive to these manipulations in predictable ways. To this purpose, the current study was successful. Coherence positively affected the accuracy of textbase links, but did not affect the accuracy of situation model links as predicted. Familiarity positively influenced both links; however, had a greater effect on situation model links as predicted.

The second goal of this study was to examine the roles of text coherence and background knowledge on the text representations of the reader. Although the first goal, stated above, indicated how coherence and background knowledge increased the rate of accuracy of both textbase and situation model connections, the second goal examined the strength of the factors that form from these connections. The third hypothesis related that the addition of macrosignals to a text would decrease the number of factors that determine the accuracy of textbase connections. Likewise, the fourth hypothesis related that the addition of familiarity of topic would decrease that number of factors that determine the accuracy of situation model connections. To this purpose, the current study was partially successful. The familiarity manipulation influenced the fit of the situational model. However, only limited evidence suggests

that the manipulation of the macrosignals increased the fit of the textbase. Following is a discussion of the findings relevant to each goal, their relations to existing research, as well as explanations on how each goal extends existing research.

PCST as a Measure of Knowledge Representation

Traditional measures of reading comprehension are predominately formed to optimize reliability and limit the number of distinct factors as well as the influence of prior knowledge (Kintsch & Kintsch, 2005; Pearson & Hamm, 2005). Although attempts have been made to update current reading measures to align with reading theory (National Assessment of Educational Progress, 2005), important information about student reading is missed by current item formats. Although the percentage correct and grade equivalent scores provided by these tests are often used to describe school wide and school district results, they are not as informative to teachers concerning the reading abilities of individual students. More promising assessments of knowledge representations built from reading comprehension are those that use proximity data to analyze the relationships readers build from concepts within a text (Ferstl & Kintsch, 1999). The test in the current study, the Passage Comprehension for Structured Text (PCST) used proximity data to measure text representations as well as components of text representations.

The first purpose of the current dissertation was to test whether these theoretical components were actually being measured by the PCST. To this end, the text manipulations of coherence and familiarity were applied to the text. The theoretical components of the PCST were the textbase and the situation model. The textbase connections measured the ability of students to identify relationships among

concepts that are stated in the text. Therefore, this information was drawn specifically from the text. The situation model connections measured the ability of students to identify non-relationships among concepts from the text. These non-relations were not stated explicitly in the text and would have to be inferred through use of comprehension as well as background knowledge. If the PCST were truly measuring the textbase and situation model components, then the components would show sensitivity to these manipulations in a way specified by reading comprehension theory as well as past reading comprehension research. Hypotheses one and two were created to test the sensitivity of the PCST measure. These hypotheses detail how, according to theory, the components of the PCST should be differentially affected by the text manipulations. The effects of text macrosignals and topic familiarity on the sensitivity of the PCST will be discussed separately.

PCST Sensitivity to the Macrosignals Manipulation

That text coherence will positively affect text recall has been well established (Britton et al., 1989; Linderholm et al., 2000). Coherence can be formed both at the local level, that is, the word-by-word or sentence-by-sentence level, and at the global level, that is, the multiple sentences or paragraph level (McNamara et al., 1996).

Macrosignals have been used to increase the coherence of the global structure of text (Loman & Mayer, 1983; Lorch & Lorch, 1996; Rickards et al., 1997). Macrosignals, as explained in the introduction, consist of “writing devices that emphasize aspects of a text’s content or structure, without adding to the content of the text” (Lorch, 1989, p. 209). These devices include titles, headings, previews, summaries, and typographical cues (Lorch, 1989, Meyer, 1975).

One goal of the current dissertation was to show the sensitivity of the PCST to the effect of macrosignals. Results of the first hypothesis indicated that students who received macrosignaled formed more textbased connections than students who read text without macrosignals. This aligns with findings from previous studies on macrosignals, which suggested that signals such as headings (Brooks, et al., 1983), previews (Spyridakis & Standal, 1987), and typographical cues (Crouse & Idstein, 1972) all have positive effects on text recall. Lorch (1989) suggested that signals would have different effects on text recall dependent on the amount of text that is signaled. Since the current dissertation study used all three types of signals, it was assumed that a large proportion of the text would be signaled. The findings verified that macrosignals did improve the formation of the textbase.

Findings of the current study also support the argument that macrosignals guide readers to use higher-order thinking strategies. One of the most salient examples of this effect was reported in a study by Mayer, Dyck, and Cook (1984). Mayer, Dyck, and Cook argued that signals in a text led readers to use a “mental model strategy” in order to form coherent representations of text. In their study one group of students was assigned a text with macrosignals and a second group was assigned a text with no signals. The no signals students were more likely to remember information in the beginning and end of the passage, which provided evidence that they processed the information in a list, whereas the signaled group remembered information throughout the passage.

Results from this dissertation indicate that the students in the macrostructure treatment formed representations that were more similar to the expert structure than

students in the no macrosignals treatment, as indicated by their correlation scores. The correlation score signifies the similarity of the student's entire representation with the expert's representation, rather than the specified components. It can be assumed from these data that the macrosignals condition led students to form more similar structures to the expert structure than the no macrosignals condition. This difference between conditions on similarity scores may indicate that the macrosignals condition was using a deeper structure strategy, such as a mental model strategy, compared to the no macrosignals condition. This extends the findings by Mayer, Dyck, and Cook (1984) by examining not just the recall of text, but the actual representations formed by the students. This argument may also explain the non-statistically significant difference found between the students in the control condition and the students in the no macrosignals condition on the measure of correlation. It is highly possible that students in the no macrosignals condition reverted to a list strategy and the control condition, having no text to read, used their structured background knowledge to assume relationships among the information. Validity of this claim could be assessed by a future study.

Although a statistically significant result was found, it is possible that an even greater difference between the macrosignal condition and the no macrosignal condition might have been observed if other changes had been made to the text. Research provides some direction as to how this difference could be enhanced. First, Lorch and Lorch (1996) suggested that the effect of macrosignals on recall increased as the complexity of the text increased. In their study, recall was assessed following a reading of a simple structured text or a complex structured text. The effects of the

macrosignals were greater when complex text was used in comparison to simpler text. The text used in the current dissertation may have been considered simple for some students with high reading ability. It could be possible for a good reader to form mental models of text without the need for signals (Mayer, Dyck, & Cook, 1980). Second, although macrosignals were shown to increase recall, especially of signaled text, this effect would be stronger if other coherence changes within the text were made simultaneously. These coherence changes include correcting coherence breaks within and between sentences, and fixing the temporal order of information within a text (Britton et al., 1989). In the current study, changes were not made to the sentences and paragraphs forming the text. It is highly possible that these other coherence changes would create an even greater difference between the two coherence conditions.

The analysis related to the first hypothesis of this dissertation also revealed that while macrosignals had a statistically significant effect on textbase connections, they did not have a statistically significant effect on situation model connections. This result supports the view that macrosignals only enhance the memory for signaled text. Evidence for this argument comes from an investigation by Lorch and Lorch (1996). In their investigation, students read a text where only half of the text information was signaled, and they were compared to students who read the same text without signals. The results indicated that while the memory for a signaled-text was improved, and the memory of a non-signaled text was recalled less. The findings of the current dissertation study, along with the findings by Lorch and Lorch, confirm that macrosignals have a differential effect on a reader's memory of text.

PCST Sensitivity to Familiarity Manipulation

Many studies have demonstrated that prior knowledge is also positively related to text recall (Britton, et al., 1998; McNamara et al., 1996; Person, Hansen, & Gordon, 1979). Readers activate prior knowledge in order to comprehend coherence gaps in text as well as form connections between the text and long-term memory (Britton, et al., 1998). Results from the second hypothesis show that familiarity had a positive influence on both textbased connections as well as situation model connections. These results are consistent with research that states that there are multiple roles of prior knowledge for building new knowledge from text.

Previous research has also examined possible interactions between familiarity and text coherence. A study by McKeown et al. (1992) indicated that coherence changes improved recall whether or not students had prior knowledge. This proved true as well in the current dissertation. Coherence changes, in the form of macrosignals, facilitated the construction of the textbase for both the students in the familiar text treatment and students in the unfamiliar text treatment. Future studies could determine whether this effect was statistically stronger in the unfamiliar text condition than in the familiar text condition. However, due to the lack of interaction among the conditions in the MANOVA of the current study, there may not be a stronger effect for the unfamiliar text condition compared to the familiar text condition.

In the same vein, it could be argued that prior knowledge is needed when a text is incoherent rather than when a text is coherent (Voss & Silfies, 1996). Voss and Silfies reported that reading comprehension ability was influential mainly when

students were reading coherent text; whereas, background knowledge was influential when students were reading incoherent text. In the current analysis, recall was not assessed separately from reading comprehension; therefore, a correlation between recall and reading comprehension could not be assessed. However, the results showed that even in the coherent text condition familiarity had influence on the scores of the students. If this result is coupled with the result of the non-significant interaction in the MANOVA, it seems that prior knowledge was needed in both text conditions.

It could also be argued that coherence will affect textbase connections; whereas, familiarity will affect both textbase and situation model connections. McNamara et al. (1996) attempted to measure knowledge built from text with short answer questions. These included textbased questions, elaborative-inference questions, bridging-inference questions, and problem-solving questions. A percentage-correct score was assigned to each item ranging from 1-2 points per item. They did find that coherent text helped students answer more textbase items correctly, but not the other three types of questions. Also, they found that background knowledge did not interact with question type. Therefore, background knowledge helped all question types evenly. These results are very similar to the results of the present dissertation study indicating that coherence influences the textbase while familiarity influences both the textbase and situation model.

However, McNamara et al. (1996) also found that coherence, especially in great amounts, is detrimental to high knowledge readers. McNamara et al. documented that high prior knowledge readers performed better on the bridging-inference questions and problem-solving questions when the text was less coherent

compared to the more coherent text. McNamara, et al. (1996) argued that the less coherent text forced the students to actively involve themselves with the reading. Highly coherent text produced a more passive reading of the text. Therefore, they suggest that high prior knowledge learners would learn better with a low coherent text.

Findings from the current dissertation do not support this assumption. In the current findings, the students in the macrosignals condition always outperformed the students in the no macrosignals condition. The discrepancy between the dissertation findings and the findings from McNamara et al. (1996) could be attributed to multiple sources. It is possible that the familiar text about America used in the dissertation study included information about America that was unfamiliar to the students. For example, the America text mentioned history facts about the Revolution. Although the information may not be new to the students, the amount of time between learning about the Revolution and the day of taking the PCST might have been so great as to affect the activation of this background information. It could also be that additional coherence changes would have made the text more coherent, and, thus, would have been closer to replicating the results found in McNamara et al. (1996). Finally, the macrosignal connections measured in the current dissertation may align more with the elaborative-inference questions in the McNamara et al. study. Elaborative-inference questions required the linking of text information with background information. These questions did not display the interaction between coherence and background knowledge.

In summary, the results from the first and second hypothesis indicated that the PCST was sensitive to the text manipulations of coherence and familiarity in the theoretically predicted manner. These findings extend measurement research several ways. First, the findings from this research signify that the use of proximity data gives a great depth of knowledge about the text representations of the students (Ferstl & Kintsch, 1999). As stated previously, important information about student reading is missed by current item formats such as multiple-choice and essay tests (Kintsch & Kintsch, 2005). The percentage correct and grade equivalent scores from these measures do not provide enough information about how students form text representations while reading (Goldsmith & Johnson, 1990; Goldsmith, Johnson, & Acton, 1991; Kintsch, 1998). Findings from this study indicate that it is possible to gain multiple scores from one PCST test that can be used to describe quite a lot about how students gain knowledge from text. First, the total correlation score indicated, in general, the similarity of the student's ratings in comparison to the expert. This score is a good measure for the general structure students build from text. Second, multiple components were assessed. The first component tested in the current study was the students' ability to form relations from concepts within the text. The ability to form coherent representations is an indicator of high-level reading comprehension ability (Gernsbacher, Varner, & Faust, 1990; Kintsch, 1998). Therefore, this is an important component to measure. The second component tested in the current study was the students' ability to use background knowledge to form inferences among concepts within the text. This second component is very similar to inference tests where the reader has to judge the validity of inferences following a short reading passage

(Hannon & Daneman, 1998). These items are similar in that the reader has to judge a non-relation, or false inference. If the reader rates these as highly similar, he did not comprehend the underlying structure intended by the author. However, unlike multiple-choice items that measure inferencing with direct questions, the PCST measures inferencing indirectly through the ratings of the conceptual relations; therefore, making it easier to construct inference items (Goldsmith & Johnson, 1990).

Although proximity data have been used successfully in past studies (Bisanz, LaPorte, Vesonder, & Voss, 1978; Britton & Gulgoz, 1991; Ferstl & Kintsch, 1999; McNamara et al., 1996; Zwaan, Langston, & Graesser, 1995), the measurement of components of proximity data has not been attempted. These results of the current study indicate that the PCST is a measure of knowledge built from text, as well as a measure of individual components of this knowledge. The components, as well as the total correlation score, not only were sensitive to the manipulations of the text, but the sensitivity was in the expected theoretical directions.

Control Analyses

Another way this dissertation expands the current literature on reading measurement is by presenting a measure that could show detailed change over time (Ferstl & Kintsch, 1999). Past research has indicated that text representations can change in structure before and after reading (Bisanz et al., 1978; McNamara et al., 1996). For example, using a numerical rating system similar to the PCST, Bisanz et al. (1978) found that students rearranged the way they organized data about animals before and after reading a story about animals. However, this change was shown through aggregated data; therefore, individual change was not examined. Likewise,

McNamara et al. (1996) used a sorting task to measure animal characteristics before and after reading a text about mammals. Students showed more awareness about mammalian characteristics in the second sorting task than in the first. The reading of the text gave the students more information about animals as well as helped the readers to see the differences between mammalian characteristics and non-mammalian characteristics. Although this showed change over time, the detailed information from the PCST would provide even more description about the changes from individual students.

Although this dissertation did not seek to measure change over time, change was simulated with a control condition. Students in the control condition were administered the two PCST forms without having read the text. The results indicated that there was a difference between the conditions for textbased connections, but not for situation model connections. These findings suggest that reading a text affected the textbased connections more than the situation connections. This would be expected since the situation model relied upon prior knowledge, whereas the textbase connections relied upon information from the text. These findings provide additional empirical evidence for the notion that textbase connections and situation model connections measure different aspects of the text representation.

This finding, although limited in scope, could lead to future studies on change over time measured with the PCST. In future investigations, measures of representations could be assessed before and after reading texts as well as before and after participating in a conceptual reading program such as Concept-Oriented Reading Instruction (Guthrie, et al., 2004).

The Salience of the Textbase and Situation Models

A second purpose of this study was to examine the roles of text coherence and background knowledge within the text representations of the reader. It has been shown extensively that both coherence and background knowledge increase text recall (Britton et al., 1989; Linderholm et al., 2000; McNamara et al., 1996; Person, Hansen, & Gordon, 1979). Further, theory suggests that both coherence and background knowledge decrease the inference load needed to comprehend the passage. While coherence decreases the need for inferences, background knowledge facilitates inference generation (Britton et al., 1998). Consequently, with coherent or familiar text, forming knowledge from text is easier and more knowledge is recalled from the text than with incoherent or unfamiliar text. This study expands on this theory by posing a further role for coherence and familiarity. In the current study, it is suggested that both coherence and background knowledge decrease the number of reader characteristics that influence how readers gain knowledge from text. Since both text characteristics decrease the load for inferencing and make it easier for the reader to gain information from text, it could be assumed that these text characteristics are, therefore, decreasing the number of factors that determine how well a student gains knowledge from a text. Stated differently, the easier the text, the fewer number of reader characteristics will affect how well the student will recall information from a text. Consequently, a stronger factor should form from the data with coherent and familiar text compared to incoherent and unfamiliar text. The current study used factor analysis to determine the strength of the factors within the different text conditions.

Effect of Macrosignals on the Saliency of the Factors

Hypothesis 3 predicted that textbase scores will be more salient and will form a more one-dimensional factor structure when students read macrosignal text compared to text with no macrosignals. If coherence of text decreases the factors necessary to form the textbase, then there should be a stronger factor when macrosignals are used compared to when no macrosignals are used. To test this hypothesis, the difference in the strength of the textbase factor in both the macrosignals and no macrosignals conditions was assessed.

To test this difference, three models were compared. These models were the homogeneous model, the partially heterogeneous model, and the fully heterogeneous model. For the comparison between the macrosignals and no macrosignals condition, the partially heterogeneous model fit the data the best. In the partially heterogeneous model the factor loadings were constrained to be the same while the residual variances and covariances were free to differ. It can be concluded, therefore, that the factor loadings of the two conditions were similar. However, the conditions did differ in the residual variances and covariances. This difference could be attributable to two sources of variation. First, the difference in residual variance could be explained by measurement error. Second, the difference in residual variance could be due to changes in other factors. Following, I will attempt to name factors that were not measured in this study, but could have related to the residual variance.

Since the results indicated that the groups were different in their residual variance and limited evidence suggests that the macrosignals condition had a higher CFI than the no macrosignals condition, it could be that there were factors that

influenced the no macrosignals condition that did not influence the macrosignals condition. These factors could be in the form of reader characteristics. As stated previously, when textbased links are signaled in the text, they become more salient in the reader's mind. The rate that the reader pays attention to the signals should mainly affect the way the student rates the textbase links. Therefore, one main reader characteristic, the attention to signals, will determine the rate that the reader will make these connections. This was partially evident in the high CFI score for the one factor model for the students in the macrosignals condition compared to the students in the no macrosignals condition. However, when no text signals are present, multiple reader characteristics will determine the amount of textbase connections that are formed. First, the reader may need to rely upon background knowledge to form the connections and fill-in coherence gaps (Britton et al., 1998). Second, reading comprehension ability, especially inference generation, would need to be relied on to make these connections (Britton et al., 1998). Lastly, other minor cuing within the text may be relied on for information about the relationships among the concepts in the textbase (Meyer et al., 1980). Therefore, when macrosignals are not used, the reader characteristics of background knowledge, reading comprehension ability, as well as attention to minor text cuing will all influence the textbase scores.

An examination of the factor loadings was also explained in the results. From these loadings, it seems that although the no macrosignals condition loadings seemed more uniform, with most loadings ranging from .42 to .50, no loadings were significant or larger than .50. In the case of the macrosignals condition, three significant loadings ranged from .44 to .62, with most falling below .37. Therefore,

the macrosignals, it seems, had more of an effect on three of the loadings. On examination, these loadings appear to be related to modern culture. Not only is the modern culture section read last, and therefore remembered better, but the section on modern culture forms a simpler structure with only one concept and three supporting details, compared to the history section.

The results indicate that the macrosignals and no macrosignals conditions were not different in their factor loadings. However, additional evidence indicates that these two conditions were different in their residual variances and covariances. Although this difference may be due to measurement error, this difference may also be due to other reader characteristics. It was suggested that background knowledge, reading comprehension ability, as well as attention to minor text cuing may be acting as the additional factors. Future studies may be needed to verify if these other factors are causing this residual variance.

Effect of Familiarity on the Salience of the Factors

Hypothesis 4 predicted that situation model scores will be more salient and will form a more one-dimensional factor structure when students read familiar text than unfamiliar text. If text familiarity decreased the factors necessary to form the situation model, then there should be a stronger factor when familiar texts are used compared to when unfamiliar texts are used. To test this hypothesis, we assessed the difference in the strength of the situation model factor in both the familiar and unfamiliar text conditions.

The results from the multiple group analysis using CFA show that the factor loadings between the two familiarity conditions were statistically different. Also,

from the individual CFA's it was shown that situation model scores from the familiar text condition had a greater overall fit compared to the scores from the unfamiliar text condition. This finding supports Hypothesis 4 that the situation model factor will become more salient and will form a more one-dimensional factor structure with the familiar text condition compared to the unfamiliar text condition.

This finding not only supports the literature on background knowledge (Britton et al., 1998; McNamara et al., 1996), it also adds information about the effects of familiarity above the evidence from Hypothesis 2. According to the results from Hypothesis 2, students in the familiar text condition performed better on the test of situation model connections than the students in the unfamiliar text condition. The findings from Hypothesis 4 add to this finding by showing that the familiar text conditions also made situation connections form a more one-dimensional factor structure than the unfamiliar text condition. Therefore, it was not just the case that one condition outperformed the other condition, but the whole nature of their ratings was changed. As expected, when the text was familiar, background knowledge becomes activated in the reader's mind. Since these connections were not highlighted in the text and the text was on a topic that was highly familiar to all the students, the rate that the reader formed accurate situation model connections depended on how well the reader activated his or her background knowledge about the topic. Therefore, one main reader characteristic, the amount of background knowledge activation, will determine the rate that the reader will make these connections. This was evident in the higher CFI score for the one factor model for the students in the familiar text condition compared to the students in the unfamiliar text condition. The higher CFI

indicated fewer factors determining the way the students formed their situation model ratings compared to the unfamiliar text condition. Although the CFI was higher than in the unfamiliar text condition, it would not be considered an excellent fit. Although activation of background knowledge would be considered a main reader characteristic for forming these connections, the activation is far from uniform. For example, one student may be more inclined to activate information about the history of America while another student may be more inclined to activate information about modern America.

Despite this limitation, the CFI for the situation model was significantly higher with the familiar text condition compared to in the unfamiliar text condition. When the text is unfamiliar, more reader characteristics determine how the students form their ratings. First, students may rely upon their background knowledge about America to answer questions about Sri Lanka. Second, when a reader is very unfamiliar with a topic he may rely upon the structure of the words rather than their meanings. For example, some students rated the words “Polonnaruwa” and “Perahera” as similar, simply because the words began with the same letter. Finally, since the PCST forces the student to make a choice, students could not skip items. Consequently many students resorted to guessing when they could not rate the item. Therefore, when unfamiliar text was presented, reader characteristics such as reliance on faulty background knowledge, word structure, as well guessing all influenced the situation model scores. This was evident in the lower CFI score for the one factor model for the students in the unfamiliar text condition compared to the students in the familiar text condition. The lower factor model suggested that more factors

determined the score pattern for the unfamiliar text condition when compared to the factor structure of the familiar text condition.

Hypothesis 3 and 4 tested the effects of macrosignals and familiarity on the factor structures of the textbase and situation model. Familiarity seemed to have a greater effect on situation model connections than the effect of coherence on textbase connections. Therefore, familiarity seems especially important to the overall structure of knowledge built from text. A greater effect of macrosignals on textbase connection might be found if more items were included. To examine whether there would be an interaction between these two attributes would be interesting. For example, would having a text that is both coherent as well as familiar produce more one-dimensional factor structures than having only one attribute present? Unfortunately, the number of students in the present dissertation was too low to perform this analysis.

The results of these analyses extend theoretical knowledge. As can be seen from the analyses of Hypothesis 4, familiarity affected the factor structures of the situation model. The situation model fit better with the students in the familiar text condition than students in the unfamiliar text condition, suggesting that familiarity of text increased the salience of these connections within the representations of the readers. Therefore, not only does familiarity of text produce more accurate ratings, these ratings form a better fitting factor. This is an additional effect of familiarity that has not been discussed in previous literature. Likewise, limited evidence suggests that text coherence is having a similar effect on the textbase. According to the strength of fit indicated by the CFI, the textbase fit better with the students in the macrosignals

condition than students in the no macrosignals condition. However, this difference may be attributable to either measurement error or other factors.

Although the use of methodologies such as factor analysis is quite common in research exploring traditional reading comprehension assessments (Davis, 1972; Pearson & Hamm, 2005) as well as other related fields such as reading motivation (Baker & Wigfield, 1999), few researchers have used this technique to examine components of reading comprehension within theoretically derived, rather than skill derived, reading comprehension measures. The results from the current study show that the use of factor analysis may be a promising methodology in studying the composition of reading comprehension tests. Through factor analyses, the additional effects of macrosignals and familiarity could be measured. Previous research has indicated that text recall increased when coherent and familiar text was used. However, through factor analysis, I was able to show that the effects of these text manipulations, especially familiarity, were more pervasive than merely increasing text recall. The number of factors determining the text representations built from text was altered. This has not been attempted in previous studies.

Additional Analyses on the Fit of Textbase and Situation Model

Additional analyses were completed to compare the strength of the two-factor structure, textbase and situation model, with the macrosignals and no macrosignals conditions, as well as the familiar and unfamiliar text conditions. Although these analyses were not particularly crucial for testing Hypotheses 3 and 4, I believed that they provide interesting findings to guide future research in this area.

By examining both the textbase and situation model scores simultaneously, the two-factor model was created. All 10 textbase items as well as all 28 situation model items were used to create this two-factor model. Figure 4 in Chapter 4 depicts the two-factor model that was tested. Both the fit of the 10 textbase items to the textbase factor as well as the fit of the 28 situation model items to the situation model factor determined the fit of the model.

The results from the multiple group analysis using CFA showed that the macrosignals and no macrosignals conditions were not different in their factor loadings when the two-factor model was fit. The best fitting 2-factor model was the partial heterogeneous model, which is similar to the result from the 1-factor model. However, there did not seem to be as great of a difference between the conditions in overall fit when the conditions were examined separately. The slight discrepancy between the fit of the one factor model and the two-factor model could be attributed to the characteristics of macrosignals. Past research has found that macrosignaled text is recalled more than non-macrosignaled text (Loman & Mayer, 1983; Lorch & Lorch, 1996; Rickards et al., 1997). Further, recall of non-macrosignaled text has been shown to decrease when macrosignals are used (Lorch & Lorch, 1996). In Hypothesis 3, I expected that the textbased factor would become more salient; however, in this expectation, I assumed that the situation model factor would remain the same for both groups. This could be problematic if the situation model had worse fit with the macrosignals condition than with the non-macrosignals condition. If this were the case, at the same time that the textbase factor was increasing in salience, the situation model factor was decreasing in salience, showing no difference in fit

compared to the non-signaled condition. The findings herein suggest that this may be the case.

The results from the multiple group analysis using CFA showed that the familiar text and unfamiliar text conditions were different in their factor loadings when a two-factor model was tested. That is, allowing the two groups to be different did statistically significantly increase the fit of the two-factor model. Again, in further analysis, findings demonstrated that situation model scores from the familiar text condition had a greater overall fit compared to the scores from the unfamiliar text condition.

In these further analyses, I found that the familiarity of text seemed to produce a larger effect on the two-factor model than the presence of macrosignals. This effect could be explained both theoretically and methodologically. Theoretically, macrosignals have been shown to have a positive influence on the recall of signaled text while simultaneously having a negative influence on non-signaled text. Therefore, as the macrosignals increased the salience of the textbase connections, they may have also decreased the salience of the situation model connections. Kintsch (1998) mentioned that as text becomes more coherent, the need for situation model connections decreases. Therefore, as macrosignals increased the coherence of text, the need for the situation model connections decreased and, therefore, the salience of the situation model connections decreased. As the textbase connections were forming a stronger factor, the situation model connections were forming a weaker structure, possibly canceling each other out. On the other hand, the results indicate that

familiarity affected the textbase and situation model both positively. Therefore, the two-factor model fit well with familiar text compared to unfamiliar text.

Methodologically, there were two times more situation model connections than textbase connections. Therefore, it would be easier to increase the fit of the two-factor structure by changing the structure of the dominate factor, the situation model, more than it would be by changing the structure of the textbase factor.

Limitations

This study used MANOVA's and CFA to examine the multivariate nature of reading comprehension. Although the results of this dissertation provide support for the multidimensionality of reading comprehension, some limitations should be taken into account when interpreting the findings.

First, the texts used in the dissertation were specifically hierarchically structured non-fictional text. Therefore, these results may not generalize to fictional stories or poetry or expository text with a different structure. For example, a similar study could occur where students are given a story about the life of a girl in Maryland or a story about a life of a girl in Tibet. I believe that, although the text structures are somewhat different, familiarity will still have an impact on both the textbase and the situation model. Likewise, I believe stories with introductory paragraphs, headings, and even captions aligned with pictures help students to form representations of the story compared to when these elements are not present. The opening to the story about the girl in Tibet could introduce the setting of the story and some background information about the girl that would help the reader to form a more coherent representation of the story. Therefore, although the text representation of a story may

be formed in a different pattern, rather than the hierarchical nature of an expository structured text, I would expect the results would generalize. However, this is not examined by the current dissertation.

A second limitation of this study was that not all components of reading comprehension were examined. In the current study, the components of textbase links and situation model links were examined. Although inferencing was assumed to be occurring for both of these components, it was not explicitly tested. In Reading Comprehension Theory, the ability to inference is a major component that determines how well a reader forms a text representation (Kintsch, 1998). Therefore, measuring the component of inference to determine how this ability influences the text representation would be important in a later study. The use of true/false statements is one measure of inference that has been used in studies (Hannon & Daneman, 1998). After reading a text section, the reader has to determine the truthfulness of statements following the reading. These statements consist of inferences that could be drawn from the text. I would expect that inferencing would impact the construction of the textbase more when macrosignals are not used compared to text where macrosignals are present. Also, I believe inferences are vital to making situation model connections.

This dissertation study was also limited by the format of the items. As described in Chapter II, there exists a multitude of reading comprehension measure formats. These include both traditional formats, such as multiple-choice and brief or extended responses, formats involving other measures of proximity data, as well as measures that have not been discussed in this dissertation such as the sentence

verification task (Pearson & Hamm, 2005). Although the components of textbase and situation model connections seem to be differentially affected by the text manipulations of familiarity and macrosignals, it may be possible that these effects were dependent on the format of the test.

For example, it would be interesting to try to isolate these components in a multiple-choice test. I believe that, although it would be difficult to form such multiple-choice items, it would be the case that these two components could be isolated within a multiple-choice test. McNamara et al. (1996) attempted to use multiple-choice, fill-in-the-blank, and short answer questions to measure components similar to those in the current dissertation. Textbase questions aligned to information presented in the text. Inference questions aligned to inferencing or analytic reasoning. Non-text questions aligned to information related to the text, but not mentioned in the text. However, they were disappointed of the results from these questions. McNamara et al. (1996) stated “The attempt to measure the deeper understanding of the problem domain that these texts communicated to the readers by means of inference questions was unsuccessful, however, perhaps because there were too few of them . . . , or perhaps we simply asked the wrong questions” (p. 14). They further mention that their sorting task, that produced proximity data, turned out to be “a more sensitive measure of situational understanding.”

Perhaps using longer essay questions could detail the students’ construction of textbase and situation models. For example, one could determine the concepts the students believe are related by how close they are associated within the essay. Although it would be difficult to objectively assign a number for the association, I

would expect that the effects of familiarity and macrosignals would be similar. As to other measures of proximity data, I believe it would be the easiest to find these effects since the format of the tests are so analogous. Future studies could address the isolation of these components within measures using other item formats.

Finally, the findings of the current study were limited to the specific students used in the study. The students in this dissertation study were all male and attended a Catholic high school. A public high school would have students of a greater variability in reading comprehension ability.

Future Directions

The findings of the current study could lead to important new research on the use of proximity data for measuring reading comprehension. One future direction for research on the PCST would be to examine other reading comprehension components, such as inferential knowledge, and to examine how these components relate to PCST component scores. Also, more components of the knowledge structure measured by the PCST could be identified.

A second future direction would be to examine the generalization of these findings with other text structures. The current study was limited to hierarchical expository texts. However, not all expository texts are hierarchical in nature. Other text structures include “description, sequence, compare–contrast, problem–solution, and causation” (Williams et al., 2005, p. 538). Since the PCST does not measure hierarchical knowledge specifically, these other text structures could be examined.

A third future direction would be to examine the generalization of these findings with other genres. Just as other text structures could be examined, other

literature genres could be used to assess comprehension. These may include fictional text such as novels and poetry. Both of these genres generally have an overall theme, main characters, and events. Therefore, a concept-map of these texts could be created and used in a PCST assessment.

A final future direction would be to examine the generalization of these findings to participants of different ages. It may be that the findings from this study may be more or less pronounced with younger students. Therefore, the isolation of the components of the knowledge structure may change as the age of the students change. This may eventually lead to developmental studies of the formation of knowledge structures.

The Birth of a Baby Shark

Some sharks can have as many as 100 babies at a time! Some baby sharks are hatched from eggs and other sharks are born live.

Sharks that lay eggs always lay them in tough cases. These cases are often called mermaids' purses. The baby grows slowly, and it may be a year old before it is ready to chew its way out of the case and swim away.

A baby shark grows in an egg case called a mermaid's purse.

Baby Lemon sharks are born live. You can see the cord attaching the baby shark to its mother just after it was born. The cord will break as the baby shark swims away. The baby shark can take care of itself right away.

Appendix B

Textbase Connections

America		Sri Lanka	
America	History	Sri Lanka	History
America	Modern	Sri Lanka	Modern
History	Settlers	History	Kingdom
History	Independence	History	Culture
Settlers	Indians	Kingdom	Polonnaruwa
Independence	Stamp Act	Culture	Perahera
Independence	Government	Culture	Kolam
Modern	Houses	Modern	Families
Modern	Food	Modern	Dishes
Modern	Job	Modern	Sari

Appendix C

Situation Model Connections

America		Sri Lanka	
History	Houses	History	Families
History	Food	History	Dishes
History	Job	History	Sari
Settlers	Modern	Kingdom	Modern
Settlers	Houses	Kingdom	Families
Settlers	Food	Kingdom	Dishes
Settlers	Job	Kingdom	Sari
Indians	Modern	Polonnaruwa	Modern
Indians	Houses	Polonnaruwa	Families
Indians	Food	Polonnaruwa	Dishes
Indians	Job	Polonnaruwa	Sari
Independence	Modern	Culture	Modern
Independence	Houses	Culture	Families
Independence	Food	Culture	Dishes
Independence	Job	Culture	Sari
Stamp Act	Modern	Perahera	Modern
Stamp Act	Houses	Perahera	Families
Stamp Act	Food	Perahera	Dishes
Stamp Act	Job	Perahera	Sari
Government	Modern	Kolam	Modern
Government	Houses	Kolam	Families
Government	Food	Kolam	Dishes
Government	Job	Kolam	Sari
Settlers	Stamp Act	Kingdom	Perahera
Settlers	Government	Kingdom	Kolam
Indians	Stamp Act	Polonnaruwa	Perahera
Indians	Government	Polonnaruwa	Kolam
Independence	Indians	Culture	Polonnaruwa

Appendix D

Method – Preliminary Investigation

In the preliminary study, I examined the differential effects of text manipulation on the text representation components of textbase and situation models. Text manipulations consisted of macrosignals and text familiarity. A control reading was also used in order to determine prior knowledge contributions to the reading assessment.

Hypotheses

In accordance with the multidimensional nature of reading comprehension, and the differential influences of text manipulations such as text familiarity and macrosignals, two research hypotheses were tested:

1. On the measure of textbase connections, students who read a text with macrosignals will outperform students who read a text without macrosignals or students who read a control text.
2. On the measure of situation model connections, students who are tested on a familiar topic will outperform students who are tested on an unfamiliar topic.

Design

This study included four one-factor designs. The first 2 one-factor designs examined the influence of the independent variable macrosignals on the dependent variables of textbase and situation model connections. Macrosignals included three levels of coherence: (a) text with macrosignals, (b) text without macrosignals, and (c) control text. The second two designs examined the influence of the independent

variable familiarity on the dependent variables of textbase and situation model connections. Text familiarity included two levels: (a) familiar topic and (b) non-familiar topic. The links used to represent the textbase and situation models were only a subset of possible links.

Participants

Participants were 60 eighth grade students. These students were from four English classes with the same teacher in a mid-Atlantic state middle school. The teacher rated her students as very below ($n = 3$), somewhat below ($n = 8$), average ($n = 32$), above average ($n = 2$), or very above average ($n = 9$) in reading comprehension. Students consisted of 45% females and 55% males.

Materials

Six texts were used in this study: (a) a familiar topic text with macrosignals on the topic of the United States of America (5 pages) that contained macrosignals such as introductions, headings, bold words, and font sizes, (b) a familiar topic text without macrosignals on the topic of the United States of America (2 pages), (c) a non-familiar topic text with macrosignals on the topic of Sri Lanka (5 pages), (d) a non-familiar topic text without macrosignals on the topic of Sri Lanka (2 pages), (e) a control text on the topic of symbiosis (4 pages), and (f) a second control text on the topic of symbiosis (4 pages). Students assigned to condition “e,” a control text, received the America PCST and students assigned to condition “f,” a second control text, received the Sri Lanka PCST. Both groups condition “e” and condition “f” received the same text.

The two texts with macrosignals were five pages in length. Most of the text and all of the pictures were pulled from two social studies trade books. The familiar text was extracted from the book *United States of America* written by Martin Hintz (2004) and the non-familiar text was extracted from the book *Sri Lanka* written by Krishnan Guruswamy (2002). Text was extracted and organized from these two books to create two texts parallel in length and organization. Both texts had one main title, two main section headings, and fifteen sub-headings. Each included seven black and white pictures. One picture in each text contained a caption. Three introductory paragraphs were written for the macrosignals texts. The first paragraph introduced the country and the text, and the next two paragraphs introduced the two main sections. The familiar reading with macrosignals had 15 paragraphs consisting of 1,012 words, and the non-familiar reading with macrosignals had 14 paragraphs consisting of 965 words. About 13 words were italicized within each text.

The two texts with no macrosignals were two pages in length. These texts were identical to the macrosignals texts; however, they did not include introductory paragraphs, headings, or italicized words. Also, these texts only included one black and white picture. The familiar reading without macrosignals had 11 paragraphs consisting of 861 words and the non-familiar reading without macrosignals had 11 paragraphs consisting of 831 words.

Measures

Passage Comprehension for Structured Text

Specifically, the current study used a measure similar to the measure used by Britton and Gulgoz (1991), which asks readers to rate the relatedness of key words

from a text. After reading their assigned text, students were instructed to rate the relatedness of 12 key words from that text, two at a time, for a total of 66 ratings.

Twelve words, representing the conceptual knowledge structure of the reading, were selected from the America text and 12 words were selected from the Sri Lanka Text. Of the 12 words, one word symbolized the main theme of the passage, three words symbolized the main concepts in the reading (two of the words representing the same concept), and eight words symbolized supporting phenomena to the two main concepts. In the America packet, *America* represented the main theme. The words representing the two concepts, found in the main headings of each section, were *early* (representing the past), *modern* (representing the present), and *everyday* (representing the present). The words *settlers*, *Indians*, *independence*, *Stamp Act*, and *government* were supporting information to past America. The words *houses*, *food*, and *job* were supporting information to the present America. These words, in hierarchical order, can be seen in Figure 5. In Figure 6 are the words selected from the Sri Lanka text.

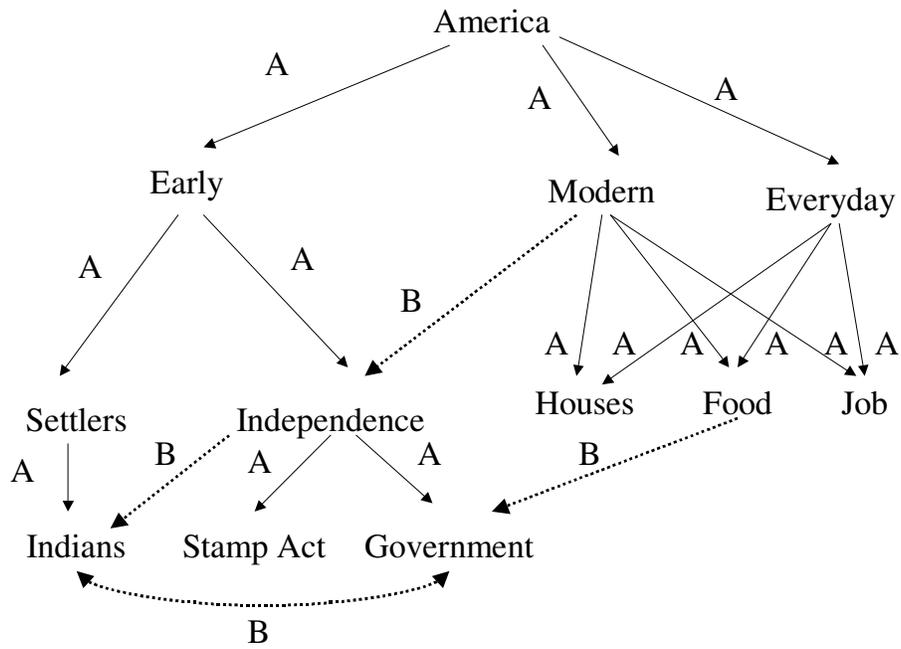


Figure 5. Representation of the text structure for America

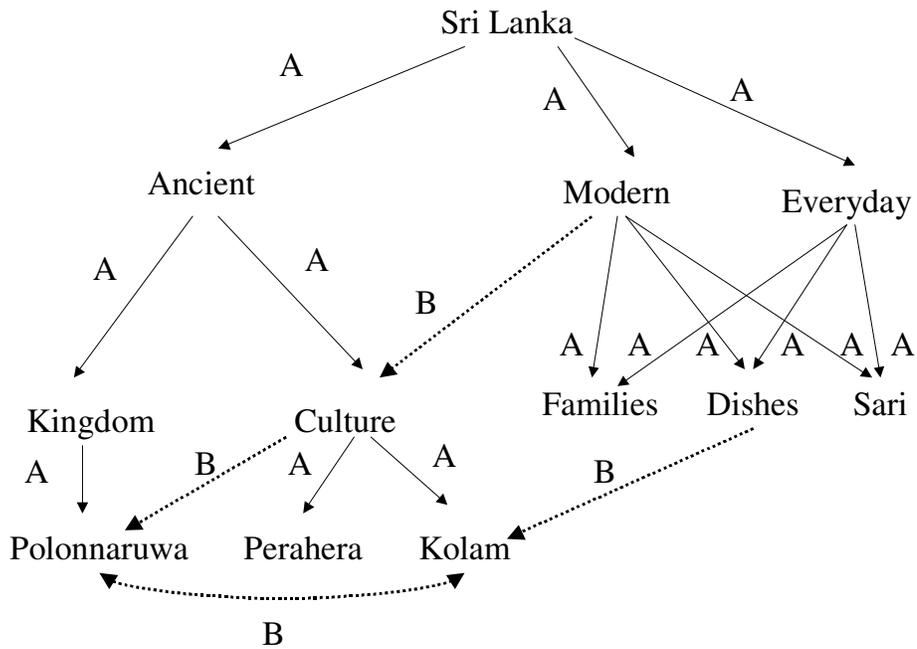


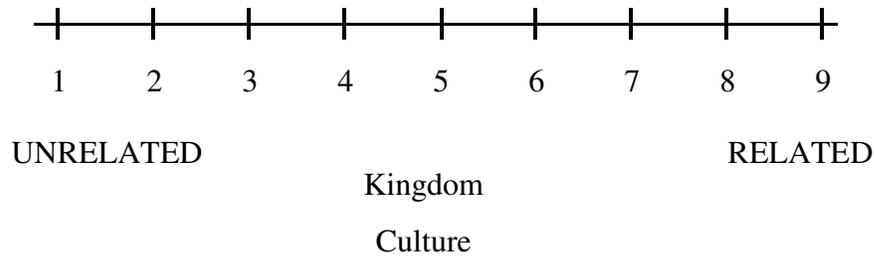
Figure 6. Representation of the text structure for Sri Lanka

An example of an expert text representation, with a few example links, can be seen in Figures 5 and 6. From the links, two comprehension components of the structure were identified as A and B. Reading comprehension component A consisted of 14 links (e.g. Links labeled A). These links are informed by textbase knowledge and are labeled on Figures 5 and 6. Component A was the sum of the links connecting the theme word with concept words (e.g. America/modern, expert rating of 9) and concept words with supporting phenomenon (e.g. modern/houses, expert rating of 9). Component B represented links involved in the distinction among concepts and phenomena (e.g. Links labeled B). These links were not explicitly stated in the text and, therefore, were informed through background knowledge. Thirty-four links were identified. Figures 5 and 6 show four sample links. Component B was the sum of the links connecting concept words with non-supporting phenomenon (e.g. modern/independence, expert rating of 1) and phenomenon with dissimilar phenomenon (e.g. food/government, expert rating of 1).

Administrative Procedures

Students were randomly assigned to read one of six texts in their classroom for 17 minutes. Following the reading passage, students were directed to the computer where they saw key words taken from the America or Sri Lanka text. Students who received the reading passages A, B, and E were administered the America PCST and students who received the reading passages C, D, and F received the Sri Lanka PCST. The computer displayed each pair of words one at a time. Students were asked to rate the words “as very related” (9), “a little bit related” (5), or “not related” (1) within the text. The original assessment was used in high school and college testing with a 1-9

scale range. Adapting it for younger students, we decided to use a three-point continuum. Figure 7 is a copy of the computer screen.



Enter: <1 through 9> followed by <SPACE>

65 Ratings to go

Figure 7. Passage comprehension computer screen

Coding Passage Comprehension Responses

Six scores were obtained from the rating data. The first score was the raw rating from each student. Although students were directed to use only the ratings 1, 5, and 9, a few students used other numbers on the continuum. In order to maintain some order, without creating a lot of missing data, raw scores were recoded such that (a) a 2 was recoded a 1, (b) a 3 was considered missing data, (c) a 4 was recoded a 5, (d) a 6 was recoded a 5, (e) a 7 was considered missing data, and (f) an 8 was recoded a 9.

The second score was the difference score and was computed by taking the difference of expert and the student rating on each pair of items. The third score was the distance score and was the absolute value of the difference score. The fourth score was the reversal of the distance score so that a high score represented high similarity to the expert and a low score represented low similarity to the expert. The fourth score was the textbase score and was computed by summing the textbase reversed

distance scores. The fifth score was the situation model score and was computed by summing the situation model reversed distance scores.

Results - Preliminary Investigation

The dependent variables examined in this study were textbase and situation model representation scores. The independent variables examined in this study were the topic familiarity (1 = familiar, 0 = not familiar), and the text type (2 = macrosignals, 1 = no macrosignals, 3 = control text). The means, standard deviations, and numbers of students in the six text conditions are presented on Table 11. The correlation matrix of these variables is shown in Table 12.

Table 11

Means and Standard Deviations of Textbase and Situation

Model Connections for Each Text

Text	Textbase	Situation Model
America Macrosignals		
Mean	48.00	157.00
SD	24.28	41.88
America No Macrosignals		
Mean	36.89	167.11
SD	20.67	32.61
America Control		
Mean	28.00	176.89
SD	20.49	45.068
Sri Lanka Macrosignals		
Mean	50.00	130.50
SD	13.18	12.27
Sri Lanka No Macrosignals		
Mean	34.22	153.78
SD	14.30	27.14
Sri Lanka Control		
Mean	41.60	146.00
SD	23.19	38.66

Table 12

Intercorrelations Among Familiarity, Macrosignals, Situation Model

Connections and Textbase Connections

	1	2	3	4
1. Familiarity	-			
2. Macrosignals	.02	-		
3. Situation Model	.33*	-.19	-	
4. Textbase	-.09	.33*	-.66**	-

Note. * $p < .05$, ** $p < .01$

Reliability Analysis

Internal constancy alphas, performed on the distance scores, were computed for the two forms (America and Sri Lanka) of the passage comprehension assessment. Distance scores were the absolute difference between the student's score and the expert score for each word-pair. The internal consistency reliabilities for the scores on the two subscales (Textbase and Situation model) of the two forms of the PCST (America and Sri Lanka) were calculated. Reliability coefficients for scores on the 14 textbase links were .87 for America and .63 for Sri Lanka. In order to increase reliability of the Sri Lanka textbase links, three links were dropped. After dropping these three links, the reliability for textbase links was .83 for America and .71 for Sri Lanka. In the following results, only the remaining 11 links were used to calculate the textbase score. Reliability coefficients for scores on the 33 situation model links were .87 for America and .62 for Sri Lanka. In order to increase reliability of the Sri Lanka situation model links, three links were dropped. After dropping these three links, the reliability for situation model links was .86 for America and .72 for Sri Lanka. In the

following results, only the remaining 30 links were used to calculate the situation model score.

Results of Analyses

Following are the results of the preliminary study organized within each hypothesis.

Hypothesis 1

On the measure of textbase connections, students who read a text with macrosignals will outperform students who read a text without macrosignals, or students who read a control text.

To examine this hypothesis, I conducted a t-test with text type as the fixed factor, and textbase formation as the dependent measure. This analysis indicated that the students who received macrosignals ($M = 50.44$, $SD = 18.36$) outperformed students without macrosignals ($M = 35.20$, $SD = 16.40$), $t(36) = 2.70$, $p = .01$, with an effect size of .40, and students who received a control text ($M = 33.80$, $SD = 22.68$), $t(36) = 2.47$, $p = .02$, with an effect size of .37. However, there was no significant difference between students without macrosignals or student with a control text, $t(38) = .22$, $p > .05$, with an effect size of .04. A box plot of these results can be seen in Figure 8. As can be seen, the macrosignals group was higher on the textbase component, indicating closeness to the expert.

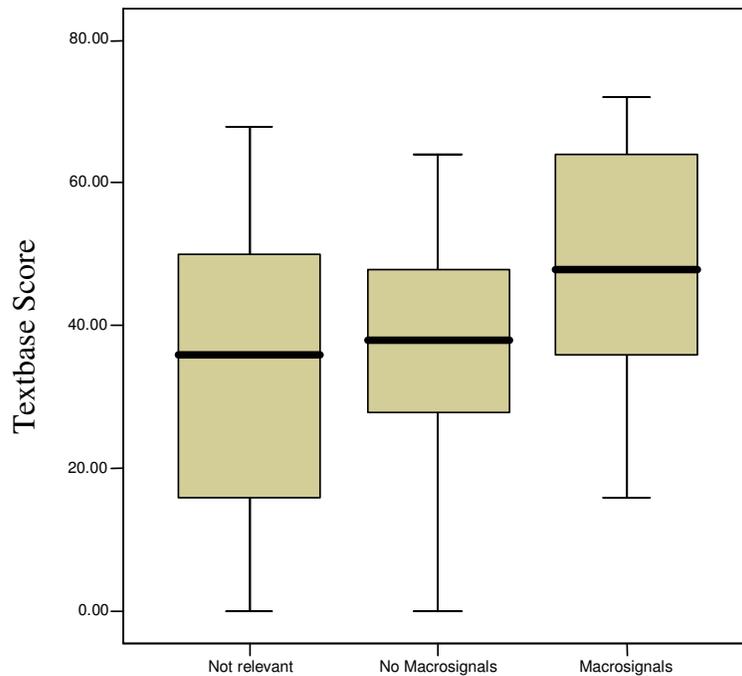


Figure 8. Student textbase scores given reading macrosignals.

A t-test with text type as the fixed factor, and situation model formation as the dependent measure was also computed. However, there was no significance between students who received macrosignals ($M = 143.75$, $SD = 32.81$) and students who did not receive macrosignals ($M = 160.44$, $SD = 29.90$), $t(32) = 1.55$, $p > .05$, with an effect size of .05, and students who received a control text ($M = 160.63$, $SD = 43.60$), $t(33) = 1.27$, $p > .05$, with an effect size of .21. A box plot of these results can be seen in Figure 9.

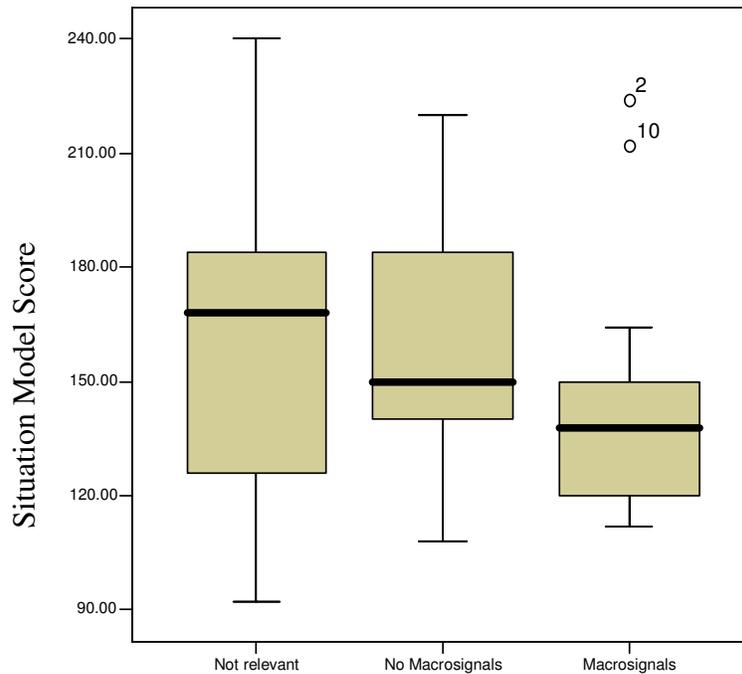


Figure 9. Student situation model scores given reading macrosignals.

Hypothesis 2

On the measure of situation model connections, students who are tested on a familiar topic will outperform students who are tested on an unfamiliar topic.

To examine this hypothesis, we conducted a t-test with topic familiarity as the fixed factor, and situation model formation as the dependent measure. Students who received the America test included both the students who read one of the America text as well as the students who read the control text and were administered the America test. Likewise, students who received the Sri Lanka test included both the students who read one of the Sri Lanka text as well as the students who read the control text and were administered the Sri Lanka test. This analysis indicated that the students who received the America test ($M = 167.38$, $SD = 39.35$) outperformed

students who received the Sri Lanka test ($M = 144.00$, $SD = 29.58$), $t(51) = 2.45$, $p = .02$, with an effect size of $.32$. A box plot of these results can be seen in Figure 10. As can be seen, the America group was higher on the situation model component than the Sri Lanka group, indicating closeness to the expert.

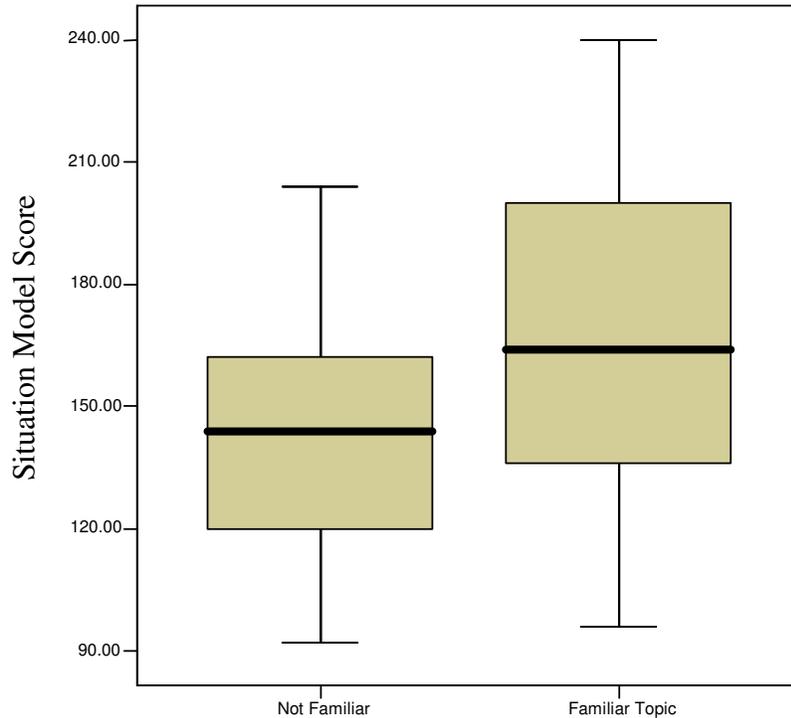


Figure 10. Student situation model scores given topic familiarity.

Next, a test between students who read an America text and students who read a Sri Lanka text was conducted. Although significance was not reached, students who read the America text ($M = 162.35$, $SD = 36.42$) did outperform students who read the Sri Lanka text ($M = 142.82$, $SD = 24.03$), $t(32) = 1.85$, $p = .07$, with an effect size of $.30$. Power analysis indicates that power is equal to $.40$. Therefore, nine subjects would be needed in each reading group in order to reach an alpha at $.05$, and 16 subjects would be needed to reach an alpha at $.01$. Unfortunately, missing data

brought the reading groups for this comparison down to about eight subjects per group.

A t-test with topic familiarity as the fixed factor, and textbase formation as the dependent measure was also computed. However, there was no significance between students who received the America test ($M = 37.73$, $SD = 22.57$) and students who received the Sri Lanka test ($M = 41.29$, $SD = 18.09$), $t(32) = .66$, $p > .05$, with an effect size of .09. A box plot of these results can be seen in Figure 11.

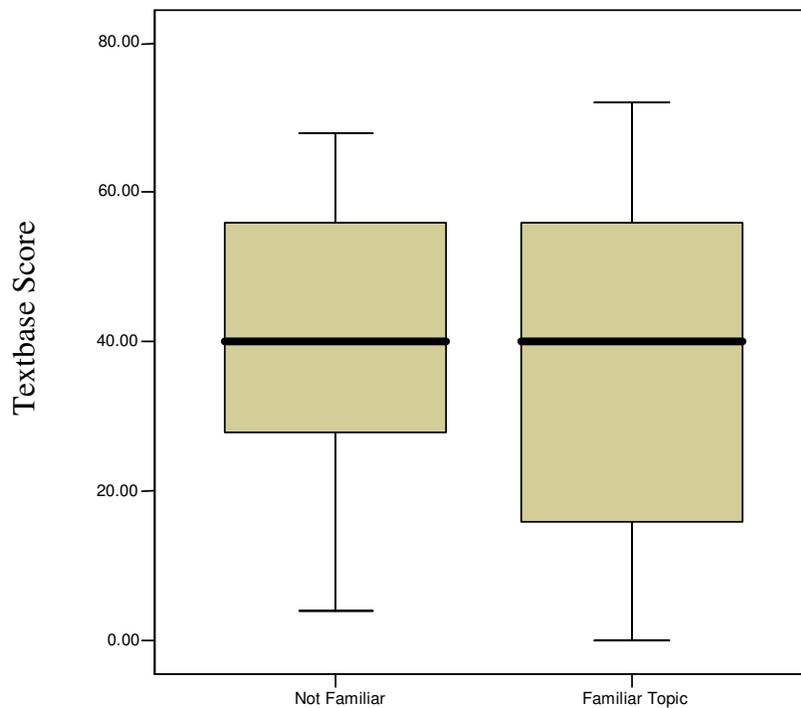


Figure 11. Student textbase scores given topic familiarity.

Discussion – Preliminary Investigation

The major purpose of this investigation was to examine the independence of the reading comprehension components of textbase and situation model connections. The first step toward understanding the independence of these variables was to

document the effects of text manipulations, specifically the addition of macrosignals such as headings, preview statements, and typographical cues, on the ability of readers to create a coherent textbase. Results of Hypothesis 1 supported this effect that has been previously explored in multiple studies on macrosignals (Loman & Mayer, 1983; Lorch & Lorch, 1996; Rickards, Fajen, Sullivan, & Gillespie, 1997). What these results add to this literature is the assessment of the textbase through a measure of the organization of the textbase. In the previous literature (McNamara et al., 1996), the textbase was measured through free-text recall and textbased questions. The textbased questions measure a reader's ability to recall information from single sentences. Although recall and textbase question measure information stored from the textbase, they do not directly measure the organization of the textbase. By using the passage comprehension assessment, one can measure individual links within the textbase, as well as the coherence of the whole textbase.

To understand further the effects of macrosignals, I examined the effect of macrosignals on situation model connections. Although the results were not significant, there was a general trend in non-signaled text and control text producing more accurate situation model connections than signaled text. Plausible explanations for the negative direction exist. For example, Lorch & Lorch (1996) found that signaled text directs a reader's attention to the information that is signaled and draws attention away from non-signaled information. Therefore, students who receive signaled text will be more likely to remember the signaled information (textbase connections) rather than non-signaled information (situation model connections).

In addition, McNamara et al. (1996) found that signaled text interfered with the formation of the situation model. In their study, high and low knowledge readers were asked to read a text with high macrostructure signaling and a text with low macrostructure signaling. Each student was asked situation model questions following reading. Students with low knowledge performed better on the situation model questions when they were given the signaling text as compared to when they were given the non-signaled text. However, students with high knowledge did better on the situation model questions when they were given the non-signaled text as compared to when they were given the signaled text. McNamara et al. (1996) believed that incoherent text forces the high knowledge reader to actively use background knowledge in order to create text representations. In view of this literature, it is believable that the signaled text would both draw attention away from non-signaled information and make the text easy enough for students not to actively use their background knowledge to form a text representation.

The second step toward accounting for the independence of the components of reading comprehension was to document the effects of familiarity on the ability of readers to create situation model connections. Results of Hypothesis 2 supported this positive effect of background knowledge on the construction of situation model connections, which has been previously explored in multiple studies on background knowledge (McNamara, 1986). What these results add to this literature is the measure of the situation model. By using the passage comprehension assessment, we can examine actual parts of the reader's text representation. This subset of situation model

connections can be compared to other connections within the representation. Also, connections can be combined to form a situation model score.

To understand further the effects of familiarity, I examined the effect of familiarity on textbase connections. Although the results were non-significant, there was a general trend in the non-familiar topic producing more accurate textbase model connections than the familiar topic. As Kintsch stated (1998) “the mental text representation is a mixture of text-derived and knowledge-derived information, not necessarily in equal parts” (p. 104). Therefore, students who have high background knowledge will have more knowledge-derived information rather than text-derived information in their final text representations. These representations will be somewhat different than textbased representations and, therefore, will not do as well in forming textbased connections. However, students with very little background knowledge, such as the students who received the Sri Lanka text, will mainly have text-derived information in their textbase and will do well making textbased connections.

The present results indicate that the two components of the text representation created through reading comprehension are differentially affected by text manipulations of coherence and familiarity. Specifically, coherent text positively affected textbase connections and marginally affected situation model connections. Further, text familiarity positively affected situation model connections and marginally affected textbase connections. This study adds to the literature by showing independence of text representation components using a measure that assesses the links in the text representation. Kintsch (1998) theorized that these two components may be inversely related, that is representations will be made of both textbase

knowledge and situation knowledge, not necessarily in equal amounts. In this way textbase knowledge and situation knowledge are inversely related. The more textbase knowledge a reader possesses, the less he or she will need to rely on situation knowledge to form a text representation. The more situation knowledge the reader possesses, the less he or she will need to rely on textbase knowledge. This study is the first to examine this relation.

The pilot investigation and the dissertation will vary in five distinct ways. First, these two studies will differ on the number of hypotheses to be examined. The dissertation will include two more hypotheses than the pilot investigation. Therefore, in the dissertation there will be two sources of information about the independence of the reading comprehension components of textbase and situation model components. The first source of information about the independence of components will derive from the text manipulations and MANOVA analyses similar to the t-tests used in the pilot investigation. The second source of information about the independence of components will derive from confirmatory factor analyses. In these analyses, the fit of a two-factor model will be collected on the textbase and situation model items. This fit will be examined for the different conditions. Second, the pilot investigation will differ from the dissertation on the macrosignals texts. Changes made to the dissertation text included adding more and different preview statements and adding more picture captions. Third, the pilot investigation will differ from the dissertation on the expert representations. One word from the each representation will be dropped and a second word in each representation will be changed. Therefore, the number of items used to form the textbase and situation model scores will be different. These

changes were informed by the pilot study investigation. Fourth, missing data will be handled differently in the dissertation than in the pilot investigation. Lastly, the pilot investigation will differ from the dissertation on the number of subjects. The dissertation will examine all of the eighth grade students in the school rather than students from one teacher.

References

- Ackerman, B. P. (1986). Referential and causal coherence in the story comprehension of children and adults. *Journal of Experimental Child Psychology, 41*(2), 336-366.
- Albright, B. F. (1927). *Typical reading disabilities of college entrants*. Unpublished master's thesis, University of Southern California.
- Andrich, D., & Godfrey, J. R. (1978). Hierarchies in the skills of Davis' Reading Comprehension Test, Form D: An empirical investigation using a latent trait model. *Reading Research Quarterly, 14*(2), 182-200.
- Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly, 34*(4), 452-477.
- Barrett, T. C. (1968). *Taxonomy of cognitive and affective dimensions of reading comprehension*. Madison (Mimeographed.).
- Barton, W. A. (1930). *Outlining as a study procedure*. New York: Bureau of Publications, Teachers College, Columbia University.
- Berry, B. T. (1931). Improving freshman reading habits. *English Journal, 20*, 824-828.
- Bisanz, G. L., LaPorte, R. E., Vesonder, G. T., & Voss, J. F. (1978). On the representation of prose: New dimensions. *Journal of Verbal Learning & Verbal Behavior, 17*(3), 337-357.

- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals [1st Edition]*. Harlow, Essex, England: Longman Group Limited.
- Britton, B. K., & Gülgöz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, 83(3), 329-345.
- Britton, B. K., Stimson, M., Stennett, B., & Gülgöz, S. (1998). Learning from instructional text: Test of an individual-differences model. *Journal of Educational Psychology*, 90(3), 476-491.
- Britton, B. K., Van Dusen, L., Gülgöz, S., & Glynn, S. M. (1989). Instructional texts rewritten by five expert teams: Revisions and retention improvements. *Journal of Educational Psychology*, 81(2), 226-239.
- Brooks, L. W., Dansereau, D. F., Spurlin, J. E., & Holley, C. D. (1983). Effects of headings on text processing. *Journal of Educational Psychology*, 75(2), 292-302.
- Brown, T. A. (2006). Confirmatory factor analysis for applied research: Methodology In *The Social Sciences*. New York: The Guilford Press.
- Cashen, V. M., & Leicht, K. L. (1970). Role of the isolation effect in a formal educational setting. *Journal of Educational Psychology*, Vol. 61(6, Pt. 1), 484-486.
- Chapman, C. A. (1969, April). *An analysis of three theories of the relationships among reading comprehension skills*. Paper presented at the meeting of the International Reading Association, Kansas City, MO.

- Cleland, D. L. (1965). A construct of comprehension. In J. A. Figurel (Ed.), *International Reading Association: Conference Proceedings, Vol. X*. Newark, Delaware.
- Clements, P. (1976). The effects of staging on recall from prose. *Dissertation Abstracts International, 37*(1-A), 190.
- Cooke, N. J. (1992). Predicting judgment time from measures of psychological proximity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(3), 640-653.
- Crouse, J. H., & Idstein, P. (1972). Effects of encoding cues on prose learning. *Journal of Educational Psychology, Vol. 63*(4), 309-313.
- Davis, F. B. (1941). *Fundamental factors of comprehension in reading*. Unpublished doctoral dissertation, Harvard University, 1941.
- Davis, F. B. (1944). Fundamental factors of comprehension in reading. *Psychometrika, 9*, 185-197.
- Davis, F. B. (1968). Research in comprehension in reading. *Reading Research Quarterly, 3*(4), 499-545.
- Davis, F. B. (1972). Psychometric research on comprehension in reading. *Reading Research Quarterly, Vol. 7*(4), 628-678.
- Dee-Lucas, D., & di Vesta, F. J. (1980). Learner-generated organizational aids: Effects on learning from text. *Journal of Educational Psychology, 72*(3), 304-311.
- Drahozal, E. C., & Hanna, G. S. (1978). Reading comprehension subscores: Pretty bottles for ordinary wine. *Journal of Reading, 21*(5) 416-420.

- Ferstl, E. C., & Kintsch, W. (1999). Learning from text: Structural knowledge assessment in the study of discourse comprehension. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading*. (pp. 247-277). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Fillmore, C. J. (1968). The case for case. In E. Bach & R. Harms (Eds.), *Universals in linguistic theory* (pp. 1-88). New York: Rinehart and Winston.
- Fillmore, C. J. (1971). Some problems for case grammar. In R. J. O'Brien (Ed.), *Report of the twenty-second round table meeting of linguistics and language studies* (pp. 35-56). Washington: Georgetown University Press.
- Frase, L. T., & Schwartz, B. J. (1979). Typographical cues that facilitate comprehension. *Journal of Educational Psychology*, 71(2), 197-206.
- Frederiksen, C. H. (1975a). Acquisition of semantic information from discourse: Effects of repeated exposures. *Journal of Verbal Learning & Verbal Behavior*, 14(2), 158-169.
- Frederiksen, C. H. (1975b). Effects of context-induced processing operations on semantic information acquired from discourse. *Cognitive Psychology*, 7(2), 139-166.
- Frederikson, C. H. (1977). Semantic processing units in understanding text. In R. O. Freedle (Ed.), *Discourse production and comprehension (Vol. 1)* (pp. 57-88). Norwood, N. J.: Ablex Publishing Corp.
- Gates, A. I. (1935). *The improvement of reading* (Rev. ed.). Oxford, England: Macmillan.

- Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3), 430-445.
- Goldsmith, T. E., & Johnson, P. J. (1990). A structural assessment of classroom learning. In R. W. Schvaneveldt (Ed.), *Pathfinder associative networks: Studies in knowledge organization*. (pp. 241-254). Westport, CT: Ablex Publishing.
- Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology*, 83(1), 88-96.
- Gonzalvo, P., Canas, J. J., & Bajo, M.-T. (1994). Structural representations in knowledge acquisition. *Journal of Educational Psychology*, 86(4), 601-616.
- Graesser, A. C., & Bertus, E. L. (1998). The construction of causal inferences while reading expository texts on science and technology. *Scientific Studies of Reading*, 2(3), 247-269.
- Gray, W. S. (1919). Principles of method in teaching reading as derived from scientific investigation. In *National Society for the Study of Education, Eighteenth Yearbook, Part II (pp.26-51)*. Bloomington, Ill: Public School Publishing Co.
- Gray, W. S. (1960). The major aspects of reading. In H. M. Robinson (Ed.), *Sequencing development of reading abilities* (pp. 8-24). Chicago: University of Chicago Press.
- Grimes, J. E. (1975). *The thread of discourse*. The Hague: Mouton.
- Guruswamy, K. (2002). *Sri Lanka*. Milwaukee, WI: Gareth Stevens.

- Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., et al. (2004). Increasing reading comprehension and engagement through Concept-Oriented Reading Instruction. *Journal of Educational Psychology*, 96(3), 403-423.
- Hand, D. J., & Taylor, C. C. (1987). *Multivariate analysis of variance and repeated measures: A practical approach for behavioural scientists*. Boca Raton, FL,: Chapman & Hall/CRC.
- Hannon, B., & Daneman, M. (1998). Facilitating knowledge-based inferences in less-skilled readers. *Contemporary Educational Psychology*, 23(2), 149-172.
- Hintz, M. (2004). *United States of America*. New York: Children's Press.
- Housner, L. D., & et al. (1993). Pedagogical knowledge structures in prospective teachers: Relationships to performance in a teaching methodology course. *Research Quarterly for Exercise and Sport*, 64(2), 167-177.
- Kintsch, W. (1974). *The representation of meaning in memory*. Oxford, England: Lawrence Erlbaum.
- Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49(4), 294-303.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.
- Kintsch, W., & Kintsch, E. (2005). Comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment*. (pp. 71-92). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

- Kintsch, W., & van Dijk, T. A. (1978). Towards a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- Linderholm, T., Everson, M. G., van den Broek, P., Mischinski, M., Crittenden, A., & Samuels, J. (2000). Effects of causal text revisions on more- and less-skilled readers' comprehension of easy and difficult texts. *Cognition and Instruction*, 18(4), 525-556.
- Loman, N. L., & Mayer, R. E. (1983). Signaling techniques that increase the understandability of expository prose. *Journal of Educational Psychology*, 75(3), 402-412.
- Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review*, 1(3), 209-234.
- Lorch, R. F., & Lorch, E. P. (1985). Topic structure representation and text recall. *Journal of Educational Psychology*, 77(2), 137-148.
- Lorch, R. F., Jr., & Lorch, E. P. (1995). Effects of organizational signals on text-processing strategies. *Journal of Educational Psychology*, 87(4), 537-544.
- Lorch, R. F., Jr., & Lorch, E. P. (1996). Effects of organizational signals on free recall of expository text. *Journal of Educational Psychology*, 88(1), 38-48.
- Lorch, R. F., Lorch, E. P., & Inman, W. E. (1993). Effects of signaling topic structure on text recall. *Journal of Educational Psychology*, 85(2), 281-290.
- Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development*, 50(3), 643-655.

- Marshall, N., & Glock, M. D. (1978). Comprehension of connected discourse: A study into the relationships between the structure of text and information recalled. *Reading Research Quarterly, 14*(1), 10-56.
- Mayer, R. E., Dyck, J. L., & Cook, L. K. (1984). Techniques that help readers build mental models from scientific text: Definitions pretraining and signaling. *Journal of Educational Psychology, 76*(6), 1089-1105.
- McKeown, M. G., Beck, I. L., Sinatra, G. M., & Loxterman, J. A. (1992). The contribution of prior knowledge and coherent text to comprehension. *Reading Research Quarterly, 27*(1), 78-93.
- McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology, 55*(1), 51-62.
- McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1-43.
- Meyer, B. J. F. (1975). *The organization of prose and its effects on memory*. New York: North-Holland Publishing Company.
- Meyer, B. J. F., Brandt, D. M. & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly, 16*, 72-103.
- Meyer, B. J. F., Middlemiss, W., Theodorou, E., Brezinski, K. L., McDougall, J., & Bartlett, B. J. (2002). Effects of structure strategy instruction delivered to

- fifth-grade children using the Internet with and without the aid of older adult tutors. *Journal of Educational Psychology*, 94(3), 486-519.
- Meyer, B. J. F., & Poon, L. W. (2001). Effects of structure strategy training and signaling on recall of text. *Journal of Educational Psychology*, 93(1), 141-159.
- Muthén, L. K., & Muthén, B. O. (2005). *Mplus user's guide. Third edition*. Los Angeles, CA: Muthén & Muthén.
- National Assessment of Educational Progress. (2005). *Reading framework for the 2009 National Assessment of Educational Progress*. Retrieved July 20, 2006, from http://www.nagb.org/pubs/reading_fw_06_05_prepub_edition.doc
- National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. U.S. Department of Health and Human Services.
- Pearson, P. D., & Hamm, D. N. (2005). The assessment of reading comprehension: A review of practices-past, present, and future. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment*. (pp. 13-69). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Pearson, P. D., Hansen, J., & Gordon, C. (1979). The effect of background knowledge on young children's comprehension of explicit and implicit information. *Journal of Reading Behavior*, 11(3), 201-209.

- Pettit, N. T., & Cockriel, I. W. (1974). Factor study of literal reading comprehension test and inferential reading comprehension test. *Journal of Reading Behavior*, 6(1), 63-75.
- RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D Program in reading comprehension*. Arlington, VA: RAND.
- Richards, I. A. (1929). *Practical criticism*. New York: Harcourt, Brace, and World.
- Rickards, J. P., & August, G. J. (1975). Generative underlining strategies in prose recall. *Journal of Educational Psychology*, 67(6), 860-865.
- Rickards, J. P., Fajen, B. R., Sullivan, J. F., & Gillespie, G. (1997). Signaling, notetaking, and field independence-dependence in text comprehension and recall. *Journal of Educational Psychology*, 89(3), 508-517.
- Rost, D. H. (1985). *Dimensionen des Leseverständnisses*. Braunschweig: Pedersen.
- Rost, D. H. (1989). Reading comprehension: Skill or skills? *Journal of Research in Reading*, 12(2), 87-113.
- Sheehan, K. M. (1997). A tree-based approach to proficiency scaling and diagnostic assessment. *Journal of Educational Measurement*, 34(4), 333-352.
- Sheehan, K. M. (2003). Tree-based Regression: A new tool for understanding cognitive skill requirements. In H. F. O'Neil & R. S. Perez (Eds.), *Technology applications in education: A learning view* (pp. 222-227). Mahwah, NJ: Lawrence Erlbaum Associates.
- Simons, H. D. (1971). Reading comprehension: Need for a new perspective. *Reading Research Quarterly*, 6(3), 338-363.

- Spearritt, D. (1972). Identification of subskills of reading comprehension by maximum likelihood factor analysis. *Reading Research Quarterly, Vol. 8(1)*, 92-111.
- Spyridakis, J. H., & Standal, T. C. (1987). Signals in expository prose: Effects on reading comprehension. *Reading Research Quarterly, 22(3)*, 285-298.
- Thorndike, E. L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology, 8(6)*, 323-332.
- Thorndike, R. L. (1973). 'Reading as reasoning.' *Reading Research Quarterly, 9(2)*, 135-147.
- Thurlow, R., & van den Broek, P. (1997). Automaticity and inference generation during reading comprehension. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 13(2)*, 165-181.
- Thurstone, L. L. (1946). Note on a reanalysis of Davis' reading tests. *Psychometrika, 11*, 185-188.
- Vidal-Abarca, E., & Sanjose, V. (1998). Levels of comprehension of scientific prose: The role of text variables. *Learning and Instruction, 8(3)*, 215-234.
- Voss, J. F., & Silfies, L. N. (1996). Learning from history text: The interaction of knowledge and comprehension skill with text structure. *Cognition and Instruction, 14(1)*, 45-68.
- Williams, J. P., Hall, K. M., Lauer, K. D., Stafford, K. B., DeSisto, L. A., & deCani, J. S. (2005). Expository text comprehension in the primary grade classroom. *Journal of Educational Psychology, 97(4)*, 538-550.

- Woloshyn, V. E., Paivio, A., & Pressley, M. (1994). Use of elaborative interrogation to help students acquire information consistent with prior knowledge and information inconsistent with prior knowledge. *Journal of Educational Psychology, 86*(1), 79-89.
- Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6*(5), 292-297.