The thesis of Melissa Ray was reviewed and approved* by the following:

Bonnie J.F. Meyer
Professor of Educational Psychology
Thesis Advisor

Robert J. Stevens
Professor of Educational Psychology

Rayne A. Sperling
Associate Professor of Education
Profession-in-Charge, Educational Psychology
Department of Educational and School Psychology and Special Education

*Signatures are on file in the Graduate School
ABSTRACT

An important aspect of expository text comprehension is structural awareness, readers’ ability to perceive and use the author’s organization to organize their own understandings. This study consisted of two experiments which examined the development of structural awareness through the comparison of 4th, 6th, 7th and 9th grade students’ performance on measures of structural awareness and text comprehension. Experiment 1 investigated the development of structural awareness for the problem-and-solution text structure, and Experiment 2 served as a replication, investigating awareness for the comparison structure. Each experiment tested for a reading skill by grade interaction on measures of text comprehension and structural awareness. Three hypotheses about the relationship between grade and skill were tested: increased difference, stable difference, and no difference. No grade by skill interaction was found. The results from both experiments support a stable difference hypothesis. Structural awareness increases with age; however there maybe periods in which little change occur.
# TABLE OF CONTENTS

LIST OF FIGURES ................................................................................................................. vi

LIST OF TABLES ................................................................................................................... vii

Chapter 1 Introduction ............................................................................................................. 1

Chapter 2 Review of Literature................................................................................................ 5

- Structural Awareness in Models of Comprehension ......................................................... 6
- Sources of Individual Differences in Structural Awareness .............................................. 10
- Comprehension Development: Implications for the Study of Text Structure ............... 16
- Focus of the Current Study .............................................................................................. 19
- Research Hypotheses ....................................................................................................... 20

Chapter 3 Methods ................................................................................................................... 23

- Experiment 1 ............................................................................................................ 23
- Experiment 2 ............................................................................................................ 31

Chapter 4 Results .................................................................................................................... 40

- Experiment 1 ............................................................................................................ 40
- Experiment 2 ............................................................................................................ 45

Chapter 5 Discussion of Results ............................................................................................. 51

References ................................................................................................................................ 56

Appendix A Problem and Solution Texts ............................................................................... 63

- Cat Text (Copyright © 2004, B.J.F. Meyer) ............................................................ 63
- Dog Text (Copyright © 2004, B.J.F. Meyer) ........................................................... 63
- Rat Text (from Meyer & Poon, 2001) ...................................................................... 64

Appendix B Problem and Solution Content Structures .......................................................... 65

- Cat Text Structure .................................................................................................... 65
- Dog Text Structure ................................................................................................... 68
- Rat Text Structure .................................................................................................... 71

Appendix C Comparison Texts............................................................................................... 75

- Penguin Text (Copyright © 2004, B.J.F. Meyer) ..................................................... 75
- Primate Text (Copyright © 2004, B.J.F. Meyer) ....................................................... 75
- Turtle Text (Copyright © 2004, B.J.F. Meyer) ........................................................ 76
LIST OF FIGURES

Figure 3-1. Incomplete Items for Experimental Tasks.............................................................38

Figure 4-1. Top Level Structure (TLS) Scores by Grade for Problem Solution Texts .............44

Figure 4-2. Competency Scores by Reading Group and Grade for Problem Solution Texts...45

Figure 4-3. Top Level Structure (TLS) Scores by Grade for Comparison Texts.....................49

Figure 4-4. Competency Scores by Reading Group and Grade for Comparison Texts...........50
LIST OF TABLES

Table 3-1. GSRT Raw, Grade Equivalent, and Quotient Scores ......................................................26
Table 3-2. Problem and Solution Recall TLS Scale (from Meyer et al., 2010).................................29
Table 3-3. Problem and Solution Recall Competency Scale (from Meyer et al., 2010)..............30
Table 3-4. Comparison Main Idea TLS Scale (adapted from Meyer et al., 2011).........................34
Table 3-5. Comparison Main Idea Competency Scale (from Meyer et al., 2010)......................36
Table 4-1. Total Ideas Recalled, TLS, and Competency for Problem Solution Texts.................42
Table 4-2. Main Ideas Recalled, TLS, and Competency for Comparison Texts .........................47
Chapter 1

Introduction

Reading expository texts can present a unique challenge to young readers. Elementary school readers may have difficulty comprehending these texts in comparison to narrative texts (Best, Floyd, & McNamara, 2008). Readers’ difficulty with expository texts likely results from a combination of prior experiences and the nature of these texts. In comparison to narrative texts, early elementary school readers may not receive the same level of exposure to expository texts (Duke, 2000). This lack of exposure could contribute to later difficulties. Additionally, expository texts may contain unfamiliar content that children, especially those with low prior knowledge, may have difficulty comprehending (Best et al., 2008).

Additionally, expository texts are organized differently from the narrative texts with which students are more familiar. While narrative texts have a focus on a series of events, expository texts have a focus on the relationship between macropropositions, the important ideas, and micropropositions, the details from the text which elaborate on them (Weaver & Kintsch, 1991). Moreover, unlike narrative structures, there is a greater variety of rhetorical structures which organize expository texts (Meyer, 1985; Weaver & Kintsch, 1991). These structures include collection, comparison, causation, description, and response (Meyer, 1975, 1985). This variety of structures coupled with a lack of exposure to expository texts may contribute to elementary and middle school readers’ challenge in understanding these texts.
Young readers are sensitive to text structure (Englert & Hiebert, 1984; Englert & Thomas, 1987; McGee, 1982; Meyer et al., 1980; Richgels, McGee, Lomax, & Sheard, 1987; Smith & Hahn, 1987; Taylor & Samuels, 1983; Yochum, 1981). However, the extent to which elementary and middle school readers are structurally aware is unclear. Several researchers have suggested that while present, structural awareness is rare in young readers. Meyer, Brandt, and Bluth (1980) found that less than 50% of students organized their recalls according to the text structure at either immediate or delayed post-tests (p. 96). Likewise, Taylor (1980) and Taylor and Samuels (1983) found that less than half of the participants in their studies could be labeled as structurally aware. This may place young readers at a disadvantage because knowledge and awareness of text structure can play a facilitative role in comprehension.

Although many students do not possess it, it is possible to acquire structural awareness. Intervention research in structure strategy training has found that with instruction, both elementary and middle school readers can gain structural knowledge and competency in using it (e.g., Armbruster, Anderson, & Ostertag, 1987; Meyer, Wijekumar, & Lin, 2011; Meyer et al., 2002; Meyer et al., 2010; Williams et al., 2005; Williams et al., 2007; Williams et al., 2009). In order to better understand the instructional needs of elementary and middle school readers, it is important to understand the role that structural awareness plays in their reading comprehension and how this role changes with age.

Previous research has shown that students’ structural awareness may vary as a function of age (Englert & Hiebert, 1984; Englert & Thomas, 1987; Garner & Gillingham, 1987; McGee, 1982; Smith & Hahn, 1989) and reading ability (Meyer et al.,
Older readers have more structural awareness than younger readers, and skilled comprehenders possess more structural awareness than less skilled readers. Less is known about the relationship between skill and age in the development of structural awareness. Research in the development of overall reading ability has generally found that comprehension skills increase with age for both skilled and less skilled readers, although differences between these groups exist and remain relatively stable (Aarnoutse, Leeuwe, Voeten, & Oud, 2001; Bast & Reistma, 1998; Shaywitz et al., 1995). Although some research findings support this relationship between age and skill in structural awareness (Englert & Hiebert, 1984; McGee, 1982), others do not (Taylor, 1980; Vauras, Kinnunen, & Kuusela, 1994). In addition to conflicting findings, a lack of research about the role of skill in the development of awareness for text structures makes drawing conclusions about the nature of this relationship difficult. Additional research is needed in order to better understand the role that reading comprehension skill plays in development.

The purpose of this study is to examine grade differences in structural awareness and whether these vary for different levels of general proficiency in reading comprehension. This will be done through the comparison of multiple grade groups’ performance on measures of reading comprehension and structural awareness. Unlike previous studies in reading comprehension development, this study will use a cross-sectional rather than a longitudinal design. Grade differences will serve as a proxy for exploring the development of awareness of expository text structures. By comparing students in several grades: fourth, sixth, seventh, and ninth, this study hopes to capture how structural awareness and comprehension may change as students progress from
elementary to high school. This study has two major goals. The first goal of this study is to extend previous research about age and skill differences in structural awareness in order to gain a clearer picture of its development. The second goal of the study is to determine if structural awareness follows similar developmental patterns as those found in previous research in reading comprehension development.
Chapter 2

Review of Literature

Structural awareness refers to a reader’s ability to recognize and use the author’s organization of a given text. The concept is derived from Meyer’s concept of a structure strategy. According to Meyer, Brandt, and Bluth (1980) readers using the structure strategy “approach text looking for patterns which will tie together the propositions contained in the text” (p.78). Within this strategy, readers’ possess knowledge of regularly occurring text structures and use this knowledge to determine the structure of the text being read and to organize their own understanding. Meyer et al. distinguished this strategy from a default/list strategy in which the text is viewed and processed as a series of loosely related propositions. Although structural awareness is frequently referred to as a structure strategy, the term structural awareness is used to here to indicate a continuum of ability ranging from no structural awareness to high levels of awareness. This is preferred for the current study’s focus on development.

The concept of structural awareness is closely tied to the concept of text structure. There are several structures including: causation, collection, comparison, response, and description which organize the macropropositions of expository text (Meyer, 1975, 1985). These rhetorical structures, also known as top-level structures (Meyer, 1985), organize those ideas from the text that are important and necessary for its comprehension. They are distinguished by the nature of the relationship between macropropositions. Collection and description are similar structures in that they include texts which are organized as a set of ideas related to the same topic (Meyer, 1975, 1985) which may or
not have a temporal relationship. Comparison refers to texts in which at least two concepts are compared and contrasted on issues. Causation refers to texts that are organized in terms of causes and effects. Finally, response refers to texts in which a problem and response to that problem, a solution, is presented. In addition to the nature of the rhetorical relationship, these structures are also distinguished by the number of constraints required by the structure (Meyer & Freedle, 1984; Meyer, 1985). Some structures, like collection and description are less restrictive in that fewer components are required to complete the structure, whereas others like causation, comparison, and response are more restrictive as more components are needed to complete the structure. Readers’ structural awareness is mapped onto these rhetorical structures, and is defined in terms of their ability to recognize and use them.

This review of literature addresses two important concerns in the study of structural awareness. The first is the role of structural knowledge in the reading comprehension process. The second focuses on factors that contribute to individual differences in structural awareness.

**Structural Awareness in Models of Comprehension**

The primary importance of structural awareness stems from its role in the comprehension process. Current models of comprehension emphasize the construction of a mental representation of text. For two models, the Construction Integration Model and the Landscape Model, this mental representation reflects the reader’s ability to understand and integrate ideas from the text into a coherent whole, the quality of which is
an indicator of comprehension. Structural awareness facilitates comprehension because it enhances the reader’s ability to construct coherent mental representations of text.

The Construction Integration Model emphasizes that comprehension and the construction of mental representations occur at multiple levels. The two primary levels are the textbase and the situational levels. These two levels are distinguished by their proximity to the text, with the textbase being closest to the actual text, and the situational level being more abstract and less connected to the text. Initially, van Dijk and Kintsch (1983) connected text propositions together on the basis of argument repetition. This was a parsimonious way to connect propositions, but insufficient for the complexity of text and the comprehension processes. The macroproposition level is more directly related to the current study and will be examined here, rather than activation and knowledge integration processes at the microproposition level. Through the use of macrostructure rules, readers can integrate the textbase with prior knowledge to construct a mental representation (Kintsch, 2004). According to the Kintsch (2004), the macrostructure rules involve the “selection” of ideas relevant to the macrostructure, the “substitution” of these specific ideas with more general (superordinate) ones, and the “construction” of macropropositions based on these superordinate ideas (p. 1282). In this process the reader identifies information important to the text and uses prior knowledge to establish macropropositions.

Similarly, the Landscape Model emphasizes the creation of a mental representation. One component of the Landscape Model focuses on levels of activation and the connections of concepts activated during the reading process. This online model of reading comprehension proposes that readers comprehend in cycles, with each cycle
corresponding to a segment of text. During each cycle, concepts or nodes are activated from the text, prior knowledge, and the “preceding processing cycle” (van den Broek, Virtue, Everson, Tzeng, & Sung, 2002, p. 137). A mental representation is not only the result of these reading cycles, it also influences these reading cycles (van den Broek et al., 2002; van den Broek, Young, Tzeng, & Linderholm, 2004). The other component central to the Landscape Model is the establishment of coherence. Coherence refers to the clarity of the connections (i.e., causal relationships) made between concepts activated during the reading process and organized in the mental representation of text (van den Broek et al., 2002; van den Broek, et al., 2004). Within the Landscape Model reading comprehension is largely a coherence seeking process.

Within both models, structural awareness can be viewed as predictor of the quality of the mental representations formed. Because structural awareness involves the reader’s ability to perceive and use the author’s organization of ideas, it can serve to improve the coherence of mental representations. Using the Construction Integration Model, structural awareness may direct the formation of the textbase and assist in the formation of macropropositions. Structural awareness can similarly be viewed in the Landscape Model. Knowledge of rhetorical relationships helps to establish strong connections between concepts. The establishment of strong, clear connections between concepts during reading can improve the coherence of the mental representation created. In both of the models, the role of structural awareness is largely top down. Like Meyer’s concept of the structure strategy, readers possess knowledge of rhetorical structures. As they read, they seek to apply these structures to the text, in order to establish its
organization, which then influences the organization of their own mental representation. The result is a well-organized mental representation.

Previous research of the relationship between text structure and reading comprehension has generally supported this role. Readers who are classified as being structurally aware recall more ideas after reading than those who are not (Meyer et al. 1980; Taylor & Samuels, 1983). Taylor and Samuels (1983) found that the benefits of structural awareness were present only when the text was well organized; for texts in which the order of ideas was scrambled there was no difference between structurally aware and unaware readers. This finding suggests that the better performance of structurally aware readers comes from their ability to perceive the macrostructure of the text. Research which has examined the effect of rhetorical form appears to support this conclusion. Kintsch and Yarborough (1982) found that readers comprehended well organized texts better than ill-structured texts, suggesting that when readers cannot determine the structure, comprehension is negatively affected. Similarly, Lorch and Lorch (1985) found that recall for scrambled text without an introduction that indicated the structure was worse than recall for well-organized texts and scrambled texts with introductions.

This effect of text structure is also present in well-constructed texts. Meyer and Freedle (1984) found that recall was greater for more restrictive structures like comparison and causation than for less restrictive structures like collection. When the structure of the text is more developed, in terms of required components, memory of the text is better. Structural knowledge can also facilitate comprehension at the local level. In this role, structural knowledge helps the reader to perceive the relationship between
sentences and phrases and as a result, can improve comprehension of these ideas. Sanders and Noordman (2000) examined the effect of listing and problem-and-solution organizations in the comprehension of sentences. The researchers found that problem-and-solution structures were read faster and recalled better than listing structures. The results suggest that when a rhetorical structure is present, sentences are processed better and faster. In general, research suggests that the ability to use structural knowledge to recognize the organization of text influences overall comprehension.

Sources of Individual Differences in Structural Awareness

Several factors predict structural awareness and should be included in studies of its development. These predictors reflect current definitions of comprehension and beliefs about sources of individual differences. According to the RAND reading study group (2002) reading comprehension is shaped by three essential elements: the reader, the text, and the activity. In this view, reading comprehension is dynamic and can vary as reader, text, and activity characteristics vary. Similarly, Meyer and Rice (1989) and Meyer (2003) asserted that the interaction among multiple variables influences the comprehension process. These variables include the reader, the task, the text, and the strategies used by the reader (Meyer & Rice, 1989; Meyer, 2003). Reader variables are characteristics unique to the reader and include age, prior knowledge, verbal ability, and working memory; text variables include text structure, topic, and genre; task variables include both how the text is presented and what readers are expected to do with the text (Meyer & Rice, 1983; Meyer & Rice, 1989; Meyer, 2003). Meyer (2003) distinguished
the strategy variables from the reader variables, referring to strategies as activities used to process a particular text.

Although all four variables contribute to structural awareness and comprehension, of primary importance in studies of young readers’ structural awareness are reader and text variables. Previous research has indicated that reader variables of age and current level of comprehension skill, as well as the text variable of rhetorical structure are significant predictors of structural awareness. Because these variables interact in reading comprehension, their contribution should be considered both individually and jointly. In addition to these variables, it is also important to consider task variables and in particular, the experimental tasks used, because differences in experimental tasks could explain discrepancies in the research literature.

Age

Central to studies of development is the relationship between age and structural awareness. Research studies which have investigated both age differences and change over time have generally found that structural awareness increases with age. Older readers recall more than younger readers and more frequently organize their recalls according to the structure of the text read (Taylor, 1980; McGee, 1982; Englert & Thomas, 1987; Smith & Hahn, 1989). In addition, older readers perform better on judgment tasks designed to measure structural awareness. In a comparison of third and sixth grade students, Englert and Hiebert (1984) found that sixth graders performed better in tasks in which readers read two stimulus sentences and were then given a set of sentences and asked to rate how each sentence fit with stimulus sentences. Sixth graders’ ratings of target sentences, those sentences which fit according to the structure of the
stimulus, were higher than third graders’, and their ratings of distractor sentences, those sentences with a different structure, were lower. Similarly, Garner and Gillingham (1987) found that when asked to organize a set of given sentences, seventh grade students scored higher on ratings of topic relatedness in comparison to fifth grade students. Moreover, a higher percentage of seventh graders demonstrated cohesion in their constructed paragraphs, than fifth graders (Garner & Gillingham, 1987).

In a longitudinal study of Dutch children’s structure use, Vauras, Kinnunen, and Kuusella (1994), measured children’s structural processing at third grade and fifth grade through analysis of recalls. Vauras et al. found that many students moved from using a list strategy to using structure in their recalls, producing recalls which used hierarchical, sequential, or associative (grouping idea on the same topic together) organization. Similar to cross sectional studies, their findings suggest that structural awareness develops with age. Moreover, their findings suggest that development marks a progression from a list strategy to a structure strategy.

**Reading Skill**

Overall reading comprehension ability is also a predictor of structural awareness and performance on measures of comprehension. Within grade groups, readers classified as highly skilled readers recall more and are more likely to use the structure of the text to organize their recall than readers classified as low skill readers (Meyer et al., 1980; McGee, 1982; Taylor, 1980). Similarly, in their study of age differences, Englert and Hiebert (1984) found that skilled readers performed better on judgment tasks of structural awareness than less skilled readers. These findings suggest that structural knowledge is associated with good comprehension. However, it is not clear whether good
comprehension leads to greater structural knowledge or structural knowledge leads to improved comprehension abilities. Intervention studies suggest the latter. Meyer et al. (2010) found that structure strategy instruction was associated with gains on a test of reading comprehension. These findings suggest gains in structural knowledge improve comprehension.

**Text Structure**

In addition to age and skill, text structure can influence structural awareness and comprehension performance. Englert and Hiebert (1984) found that performance on rating tasks was related to text structure for both third and sixth graders. Both sixth and third graders performed worse on rating items from comparison and description structures than on those from enumeration and sequence structures (Englert & Hiebert, 1984). Moreover, they found an interaction between grade and text structure on ratings of distractor items. Third grade students performed better with sequence than description, while there were no significant differences in ratings across text types for sixth graders (Englert & Hiebert, 1984). Similarly, Smith and Hahn (1989) found that in relation to enumeration and description, comparison was used the least often to organize recall. They also found a relationship between age and structure. All fourth, sixth, and eighth grade students used description and enumeration structures to organize oral recalls; however, for the comparison structure few fourth and sixth graders used the structure, while many eighth grade readers did (Smith & Hahn, 1989). These researchers’ findings suggest that structural awareness varies by text type; readers may have greater levels of awareness for some structures than others. This research also suggests that text structures
may develop differentially, with awareness for some structures being acquired before others.

Although there is some evidence that structural awareness may vary as function of text, it is not clear which structures are easier and which are more challenging for young readers. Richgels, McGee, Lomax, and Sheard (1987) also found that certain text structures were associated with higher amounts of structural awareness. Unlike Englert and Hiebert (1984) and Smith and Hahn (1989), Richgels et al. found that use of structure was greatest for comparison texts, in relation to enumeration, description, problem-and-solution, and causation structures. Readers demonstrated the least amount of structural awareness for the causation structure; low performance in using the causation structure was found for both recall and compositions tasks (Richgel et al., 1987). The discrepancy in findings about the difficulty of comparison could be due to differences in experimental tasks and texts. It also might result from the specific types of structures being compared. While Englert and Hiebert and Smith and Hahn compared enumeration, description, sequence, and comparison, Richgels et al. compared enumeration, comparison, causation, and problem-and-solution. Despite this discrepancy, previous research does indicate that studies of development should consider the influence of text structure on children’s performance on measures of structural awareness.

**The Relationship between Age and Skill**

Although age, skill, and text structure predict structural awareness, it is the relationship between age and skill that is most important for studies of development. A consideration of both age and skill can help explain why some readers possess structural awareness while others do not. Several studies which have examined age differences also
have accounted for the role of skill. Taylor (1980) compared adults, highly skilled sixth grader readers, low skill sixth grader readers, and skilled fourth grade readers in the use of structure and the number of ideas recalled. Taylor found that at immediate posttest, sixth graders recalled more ideas than fourth graders, with no significant differences between skilled and less skilled readers. At delayed recall, skilled sixth grade readers recalled more ideas than both less skilled sixth grade and fourth grader readers with no significant difference between the less skilled sixth graders and the fourth graders (Taylor, 1980). Similar results were found for the structure of recalls. At immediate recall, there were no differences between these three groups; however, at delayed recall a greater number of skilled sixth grade readers used the text structure in comparison to the less skilled sixth graders and the fourth graders (Taylor, 1980). No significant differences were found between fourth graders and less skilled sixth graders. These findings suggest that an examination of age differences should also include reading ability because age differences may be contingent upon reading skill.

Vauras et al. (1994) longitudinal study also supports this idea. Vauras et al. compared the growth of children classified into three ability groups: high, middle, and low. The researchers found that use of structure in recalls significantly improved for the high and middle groups, but that for the low group the difference between third and fifth grade performance was not significant. Moreover, both local and global cohesion ratings improved for the high and middle groups, but only local coherence improved for the low group (Vauras et al., 1994). Vauras et al. found that the greatest amount of change occurred in those readers classified as high ability and lowest with those classified as low ability. Like Taylor (1980), these findings suggest that age differences are related to the
skill of the reader. In addition, the findings suggest that reading skill can influence the extent to which structural awareness develops with age.

Other research suggests that although age and skill predict structural awareness, reading skill does not influence age differences in structural awareness. In a study similar to Taylor (1980), McGee (1982) compared the structural awareness of high ability third grade readers to high and low ability fifth grade readers. McGee found that high ability fifth graders recalled more than the low ability fifth graders and the third graders. Unlike Taylor, McGee found that all fifth graders recalled more than third graders. The students’ use of structure followed a similar pattern. The high fifth-grade group mostly used the full structure, the low fifth-grade group demonstrated partial use of the structure, and the third grade group showed no structure (McGee, 1982). In contrast to Taylor, McGee found that even less skilled older readers performed better than skilled younger readers. Englert and Hiebert (1984) found similar results. They compared the performance of third and sixth grade students who were classified into one of three reading ability groups: low, medium, and high. They found main effects for both age and skill, but no significant skill by age interaction. This finding suggests that age differences are not moderated by reading skill.

**Comprehension Development: Implications for the Study of Text Structure**

In addition to considering predictors of structural awareness, studies of its development also need to consider individual differences in general comprehension development. Specifically, studies should consider the relationship between age and skill in overall comprehension. This is important for two reasons. First, because structural
awareness can be considered an important component of reading comprehension, it is necessary to consider whether the development of structural knowledge follows a similar pattern to the development of general comprehension ability. Second, because there is little research in the development of structural knowledge, previous research in general reading comprehension can provide an additional means of framing research questions related to development.

Perspectives on reading comprehension development provide contrasting explanations of the relationship between ability and age. These views are similar to those found in research of structural awareness development. One view, the Mathew effect (Stanovich, 1986), suggests that as readers age, the difference between skilled and less skilled readers increases. According to Stanovich, there is reciprocal relationship between abilities like vocabulary and comprehension, such that readers who have good vocabularies are able to use these abilities to improve comprehension, and comprehension in turn improves vocabulary. As a result of this reciprocal relationship, good readers experience a higher quantity and quality of reading experiences than poor readers, which in turn leads to greater comprehension differences (Stanovich, 1986).

There is some empirical evidence to support this theory. In a longitudinal study of good and poor readers, Juel (1988) found that from first through fourth grade, readers stayed in their ability groups, and that while good readers’ reading and writing abilities improved, poor readers failed to make these gains. However, it is not clear from Juel’s findings whether reading comprehension differences increased overtime. Juel suggested that poor decoding abilities contributed to poor vocabulary growth, exposure to print, and a dislike of reading, which in turn may have contributed to little gain among poor readers. In other
words, the inability to master lower level reading processes precluded the development of higher level reading processes like comprehension.

The alternative view suggests that differences between skilled and less skilled readers do not increase with age. In this stability view of reading development, both skilled and poor readers’ comprehension abilities increase over time; however, the differences between the groups do not change significantly. Previous research in reading development supports this view. Shaywitz et al. (1995) found that from kindergarten through 6th grade, children remained in their comprehension ability group, such that low students stayed low and high students stayed high. Unlike the Mathew effect would suggest, the differences between the high and low groups did not significantly increase over time. Moreover, the rate of growth for students in the low group was greater than students in the high group, even though overall their reading comprehension was significantly worse (Shaywitz et al. 1995). Similarly, in a longitudinal study of Dutch children, Aarnoutse, Leeuwe, Voeten, and Oud (2001) found that differences in comprehension between high, middle, and low reading ability groups remained stable over time, and the effects sizes for the low group were greater than the high and middle groups. These findings suggest that poor readers do make gains, but because they begin the school year with low reading abilities, they are unable to “catch up” to average and high readers.

Both perspectives indicate that studies of development should account for the role of reading skill. Within the Mathew effect, prior reading ability can actually constrain reading development, leading poor readers to fail to acquire the reading skills that good readers do. For the stability view, reading skill is also powerful predictor of reading
comprehension. However, for this view, reading skill is a better predictor of within grade than across grade differences. Given the presence of these conflicting viewpoints in both general comprehension and structural awareness research, studies of development should address these two theories about the relationship between age and skill.

**Focus of the Current Study**

The current study expands on previous research in structural awareness by examining the relationship between age and skill in structural awareness of multiple text structures. Previous research in awareness of text structure has typically examined one age group (Meyer et al., 1980; Richgels et al., 1987) or two grades (McGee, 1982; Englert & Hiebert, 1984; Englert & Thomas, 1987; Garner & Gillingham, 1987). As a result, it provides a limited description of age differences. This study expands the number of age groups under examination from two to four in order to gain a clearer picture of age differences. In addition, although previous research has examined several different text structures, some text structures like description, collection, and comparison are frequently studied, while others like causation and problem-and-solution are less frequently studied. Although comparison has been frequently studied, research has provided somewhat conflicting findings regarding its awareness in relation to other structures. Although the current study does not present a direct comparison of text structures, it contributes to the field through the exploration of awareness of a less studied structure (problem-and-solution) and a structure in which competing findings exist (comparison).

This study is comprised of two experiments, which will answer the same research questions:
1) Does structural awareness increase with age and reading skill?

2) Is there a skill by age interaction on structural awareness measures? In other words, do students’ reading comprehension abilities influence age differences in structural awareness?

3) Does comprehension of expository text increase with age and reading skill?

4) Is there a skill by age interaction on comprehension performance? In other words, do students’ reading abilities influence age difference in text comprehension?

Experiment 1 will examine structural awareness for problem-and-solution texts and Experiment 2 will examine structural awareness for comparison texts. Experiment 2 is designed to test the replication of findings in Experiment 1, but with another text structure. The two structures are not being directly compared, because students completed different tasks for each structure (recall for Experiment 1 and main idea task for Experiment 2), which does not allow for direct comparison.

**Research Hypotheses**

Both Experiment 1 and Experiment 2 will test three hypotheses about the relationship between grade and skill for structural awareness and comprehension. Two hypotheses are derived from previous longitudinal research in reading development. The first of which is the increased difference hypothesis. Akin to the Mathew effect (Stanovich, 1986), this hypothesis asserts that skill differences will increase across grade groups. The second hypothesis is the stable difference hypothesis, which suggests that significant differences in reading skill groups will be present at each grade, but these
differences will not be significantly different from each other. The third hypothesis, the *no difference hypothesis*, is derived from previous research in awareness of particular rhetorical structures.

Studies which have examined the performance of children have found that certain structures are more difficult to acquire than others. Although there are conflicting findings regarding which structures are more challenging, one consistency in the results is that text structures that have a constrained structure (a number of required component parts) and require abstraction of concepts are associated with poorer performance than structures with temporal relations (sequence) or non-restrictive structures (enumeration) in which little abstraction is required (Englert & Hiebert, 1983; Richgels et al., 1987).

In contrast, research that has examined adults’ awareness of structures (Meyer & Freedle, 1984; Sanders & Noordman, 2000) suggests that the stronger the contingent relationships, and the more restrictive the structure, the greater the facilitation in comprehension. This seemingly contrasting relationship between adults and children suggests that while causation, problem-and-solution structures (which inherently contain casuals), and comparison may be helpful for competent, adult readers, younger readers may still be developing knowledge and use of these structures. This development may reflect the challenge younger readers face when comprehending expository text. Thus, the *no difference hypothesis* suggests awareness of these structures and competency in comprehension may only be present in the oldest, most skilled readers. According to the *no difference hypothesis*, structural awareness requires both age and skill to develop and as a result, there will be no significant differences among all groups with the exception
that the most skilled, oldest readers (9th grade) will perform better on structural awareness and comprehension measures in comparison to all other grade and skill groups.

It is hypothesized that although there will be significant differences across grade groups, and within each grade group, among skilled and less skilled readers, there will be no interaction between grade and skill. Thus, results from the current study are predicted to support the stable difference hypothesis.
Chapter 3

Methods

Experiment 1

Design

The current study is an analysis of data collected as part of previous study on children’s and young adolescents’ knowledge of text structure. The previous study was a pilot study to examine the use of texts and tasks to measure students’ comprehension and structural knowledge. This study is a descriptive study. The independent variables are age, which is defined as the students’ grade, and reading ability, which is defined by students’ performance on a standardized reading test. The dependent variables are ideas recalled, top-level-structure score, and structure competency score.

Participants

The participants were students from grades 4, 6, 7, and 9 from same the school district, which included three elementary schools, one middle school, and a high school. The school district is located in a small city in western Pennsylvania. In this particular district, the middle school and high school are served by one school campus, and the elementary schools are comprised of grades K-6. There were 498 participants across all four grades. In order to be included in this analysis a score on a standardized test of reading comprehension was needed. Students for whom there were no reading comprehension test scores were excluded from further analysis. Additionally, students who did not complete any tasks in the packet of materials, or only attempted the first task were excluded from analysis. After the exclusion of these students, the sample was 437
students with 123 fourth graders, 160 sixth graders, 73 seventh graders, and 81 ninth grade students. Of those students excluded from the analysis, six were fourth graders, 13 were sixth graders, 30 were seventh graders, and 12 were ninth graders.

In each grade, students were assigned to one of three reading ability groups: low, middle, and highly skilled readers based on standardized reading comprehension test scores (Gray Silent Reading Test [GSRT]; Wiederholt & Blalock, 2000). This test is designed for both group and individual administration. The administration and scoring of the test are based on basals and ceilings. Stories in the test are ordered by difficulty with five multiple choice questions per story. For group administration, readers are given a set of stories appropriate for their age and basal and ceilings are assumed by test performance (all five questions per story answered correctly yields a student’s basal score; three or more incorrect answers for a passage signals the ceiling and no points are accumulated for that story or later stories) (Wiederholt & Blalock, 2000).

Within each grade, students were assigned into one of three reading skill groups (low, middle, high) using scores from the GSRT. Students’ raw scores were converted into z- scores. Students with z-scores of 0.55 or greater were assigned to the high group, students with scores between -0.55 and 0.55 were classified as middle, and students with z-scores of -0.55 and below were classified as low. Additionally, grade equivalent scores and quotient scores were determined for each student. Quotient scores are standard scores calculated with the raw scores and the child’s chronological age. These scores have a mean of 100 and standard deviation 15 (Wiederholt & Blalock, 2000, p. 27). The benefit of quotient scores is that students’ performances are compared to a normed sample rather than just the children in the school. Quotient scores were not used to classify students
because they are based on chronological age, and these ages may or may not coincide with the student’s grade. Table 3-1 contains GSRT raw, grade equivalent, and quotient scores.

Using z-scores, the following samples sizes for each group were determined. For fourth grade there were 36 in the high group, 54 in the middle group, and 33 in the low group (Raw Score $M = 25.21$, $SD = 11.33$). For sixth grade there were 58 in the high group, 50 in the average group, and 52 in the low group ($M = 34.16$, $SD = 13.44$). For seventh grade there were 23 in the high group, 29 in the average group, and 21 in the low group ($M = 35.93$, $SD = 10.27$). For ninth grade there were 21 in the high group, 42 in the average group, and 18 in the low group ($M = 43.65$, $SD = 13.02$). An ANOVA conducted on raw scores, with grade group as the independent variable, showed a main effect for grade $F(3, 433) = 38.02$, $p < .0005$. A post hoc comparison of grades using Tukey HSD indicated significant differences between all grades with the exception of the sixth and seventh grade groups; the difference between the sixth grade ($M = 34.163$, $SD = .973$) and the seventh grade ($M = 35.932$, $SD = 1.44$) mean raw GSRT scores was not significant, $p = .309$. 
### Table 3-1. GSRT Raw, Grade Equivalent, and Quotient Scores

<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>Raw Score</th>
<th>G.E.</th>
<th>Quotient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Female</td>
<td>Male M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Grade 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>33</td>
<td>16</td>
<td>16</td>
<td>12.36</td>
</tr>
<tr>
<td>2</td>
<td>54</td>
<td>25</td>
<td>26</td>
<td>23.61</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>18</td>
<td>17</td>
<td>39.42</td>
</tr>
<tr>
<td>Grade 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>52</td>
<td>12</td>
<td>34</td>
<td>18.06</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>20</td>
<td>24</td>
<td>35.00</td>
</tr>
<tr>
<td>3</td>
<td>58</td>
<td>24</td>
<td>25</td>
<td>47.88</td>
</tr>
<tr>
<td>Grade 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>21</td>
<td>11</td>
<td>9</td>
<td>22.62</td>
</tr>
<tr>
<td>2</td>
<td>29</td>
<td>23</td>
<td>3</td>
<td>36.72</td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>6</td>
<td>15</td>
<td>47.09</td>
</tr>
<tr>
<td>Grade 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>18</td>
<td>5</td>
<td>11</td>
<td>23.44</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td>20</td>
<td>17</td>
<td>45.71</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>10</td>
<td>10</td>
<td>56.86</td>
</tr>
</tbody>
</table>

*Note. 1= low reading group, 2= middle reading group, 3= high reading group*
Materials/ Data Sources

**Gray Silent Reading Test (GSRT).** Reading ability was assessed using the Gray Silent Reading Test (GSRT) which is a multiple choice, standardized reading test. The test is designed to help identify students in need of additional assistance, and because it is sensitive to instruction, to determine growth in comprehension (Wiederholt & Blalock, 2000, p.6). The reliability for the form used in this study (Form A) is .95.

**Texts.** There were three texts: cats, dogs, and rats. These texts were adapted from news articles and designed to be equivalent (see Appendix A). According to the Flesch-Kincaid grade level readability index, the readability of each text is approximately equal (rat = 10.5, cat = 10.8, dog = 10.4). A content structure analysis was performed for each text to determine the organization of the text and the number of propositions for each text, using Meyer’s (1975, 1985) method of analysis. The structures used in this study were created by the authors of the pilot study from which the data were taken. Each text had the same top level structure, problem-and-solution (response). Each text also had the same number of proposition units with a total of 72 units (see Appendix B).

**Tasks.** For each problem-and-solution text, students were instructed to read the text. After reading, students removed the text from the testing packet and placed it into an envelope. Students were then asked to write everything that they could remember about the text.

**Procedures**

All of the texts and tasks for Experiments 1 and 2 were contained in the same a packet of written materials. Prior to completion of the packet, students completed the GSRT. Students completed all study items in one session, individually. The problem-
and-solution texts in the packets were counterbalanced. There were three possible orders (C [cats], D [dogs], R [rats]; D, R, C; R, C, D). Problem-and-solution texts were interspersed with comparison texts, such that between each problem-and-solution text, students read comparison texts.

**Scoring**

**Total Recall.** Using the content structure of each text, the total number of ideas recalled from each text was calculated for each student. Each participant received a total recall score which was the total number of ideas summed across all texts. For each text for which no recall was present, students received a score of 0 ideas recalled. The scoring from the original pilot study was used. Ten percent of the sample was randomly selected, and these students’ recalls were also scored by the author. The percentage of rater agreement was 0.99.

**Top-level-structure (TLS) score.** Using Meyer and colleagues’ (Meyer et al., 1980; Meyer, 1985; Meyer et al., 2002; Meyer et al., 2010; Meyer, Wijekumar, & Lin, 2011) top-level-structure scale, each recall was given a top level structure score. Top-level-structure scores are a measure of students’ organization of ideas according of the author’s rhetorical structure. This score is designed to reflect the organization separate from the correctness of the content of the recall. TLS scores range from 1 (no use of the structure) to 9 (use of the structure with signaling of the structure) (see Table 3-2). Scores 6-9 indicate an awareness of the problem-and-solution text structure. Students received a TLS score for each recall, and then a mean TLS score was calculated for each student. For each text read by a student that yielded no ideas remembered, students received a TLS score of 0. All recalls were scored by the author. Ten percent of the
sample was randomly selected and these students’ recalls were scored by a second rater.  

The percentage of rater agreement was 0.86.

Table 3-2. Problem and Solution Recall TLS Scale (from Meyer et al., 2010)

<table>
<thead>
<tr>
<th>Score</th>
<th>Description of TLS Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No correspondence (“I don’t know,” “I don’t remember”) or 90% or more of ideas are not from the article.</td>
</tr>
<tr>
<td>2</td>
<td>Descriptive list of ideas about the text with no indication in any sentences about text structure used as the top-level structure of the article. (“In the story I remember seeing these words ‘National Institute of Health.’ Psychologists are working with rats.”)</td>
</tr>
<tr>
<td>3</td>
<td>More than a descriptive list (use of other structures, such as cause-and-effect) with no ideas organized as a solution to the problem.</td>
</tr>
<tr>
<td>4</td>
<td>The student’s recall is organized like a list of ideas remembered. Like 2 above, but within one of the listed ideas (within a sentence or 2 adjacent sentences) is organized with a problem and solution, the structure of the article.</td>
</tr>
<tr>
<td>5</td>
<td>Like 3 above, but within a sentence or two adjacent sentences student expresses the same text structure as that used to organize the article (problem &amp; solution)</td>
</tr>
<tr>
<td>6</td>
<td>Top-level structure of recall matches that of the article, but no explicit signaling. (All the problem ideas are presented together followed by all the solution information – even if the content of the problems or solutions do not match that of the text.)</td>
</tr>
<tr>
<td>7</td>
<td>Top-level structure of recall matches that of the article, and explicit signaling of first part of the text structure. (Problem signaling word explicitly cues problem part.)</td>
</tr>
<tr>
<td>8</td>
<td>Top-level structure of recall matches that of the article, and explicit signaling of second part of the text structure. (Solution signaling word explicitly cues solution part.)</td>
</tr>
<tr>
<td>9</td>
<td>Top-level structure of recall matches that of the article, and explicit signaling of both parts of text structure. (Both problem and solution signaling words used in recall.)</td>
</tr>
</tbody>
</table>
Competency in structure scores. Using Meyer et al.’s (2010) competency scale, each recall was given a competency score. The competency scale is distinct from the TLS scale because it measures both organization and correctness of remembered ideas. As result, it is an indicator of both comprehension and structural awareness. Competency scores range from a score of 1 (no use of the structure), to 6 (use of the structure) (see Table 3-3). Scores of 5 and 6 indicate identification of the correct problem and the correct solution. Students received a competency score for each recall, and then a mean competency score was calculated for each student. For each text for which no recall was present, students received a competency score of 0. All recalls were scored by the author. Ten percent of the sample was randomly selected and these students’ recalls were scored by a second rater. The percentage of rater agreement was 0.81.

Table 3-3. Problem and Solution Recall Competency Scale (from Meyer et al., 2010)

<table>
<thead>
<tr>
<th>Score</th>
<th>Description of Competency Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Recall contains no problem, no cause, and no solution.</td>
</tr>
<tr>
<td>2</td>
<td>Recall contains a signaled cause, but no problem and no solution. Content information about the problem may be mentioned, but problem is not developed.</td>
</tr>
<tr>
<td>3</td>
<td>Recall contains the problem, but no solution &amp; no cause. Alternatively, recall presents a solution without a problem and cause.</td>
</tr>
<tr>
<td>4</td>
<td>Recall contains the problem(s) with cause(s), but no solution(s). Alternatively, the recall has the correct problem(s) with its correct cause(s), but the solution is incorrect. Signaling is not required for problem</td>
</tr>
<tr>
<td>5</td>
<td>Recall contains the correct problem(s) and the correct solution(s) but no cause.</td>
</tr>
<tr>
<td>6</td>
<td>Recall contains the correct problem(s) with cause(s), and the correct solution(s) that can help to eliminate the cause(s).</td>
</tr>
</tbody>
</table>
Experiment 2

The design and participants for Experiment 2 were the same as Experiment 1.

Materials/Data Sources

Texts. There were six comparison texts: turtles, primates, penguins, patriots vs. loyalist women, continental vs. British army, and Mount Rushmore vs. Easter Island (see Appendix C). Three of the texts were experimenter controlled (turtles, primates, and penguins) and designed to be equivalent. The readability of each text was assessed using the Flesch-Kincaid grade level index. The readability levels of the experimenter controlled texts were similar: penguin = 6.1, turtles = 5.5, primates = 5.4. The readability of the other passages was somewhat higher: patriots vs. loyalist women = 7.8, continental vs. British army = 9.0, Mount Rushmore vs. Easter Island = 7.1. The historical texts were taken from a fifth-grade history textbook (Berson, 2003, p. 311 & 315). The Mount Rushmore vs. Easter Island text was written after examining literature about the construction of stone heads at the two sites, a topic used successfully in a study about text structure and writing with eighth-grade students (Hammann & Stevens, 2003). A content structure analysis was performed for each text to determine the organization of the text and the number of propositions for each text, using Meyer’s (1975, 1985) method of analysis. The structures used in this study were created by the authors of the pilot study from which the data was taken. Each text has the same top level structure, comparison. The experimenter controlled texts have the same number of propositional units with a total of 91 units (see Appendix D). The other passages varied in total numbers of ideas: the patriots vs. loyalist women text had 97 total ideas, the continental vs. British army
text had 117 total ideas, and the Mount Rushmore vs. Easter Island text had 182 total ideas.

**Tasks.** For each comparison text, students first completed the passages with missing comparison signaling words, and then wrote a main idea for the passage they had just read. Students were instructed to write no more than two sentences. The text remained available to the students during the writing task. The signaling task was not analyzed in this study.

**Procedures**

The procedures for Experiment 2 are the same as Experiment 1, with comparison texts interspersed with problem-and-solution texts. The experimenter controlled, equivalent texts (turtles, penguins, primates) were counterbalanced. These texts appeared as either the first, fourth, or sixth comparison text in the packet. There were three possible orders of comparison texts (Penguins, Turtles, Primates; Primates, Penguins, Turtles; Turtles, Primates, Penguins). These corresponded to the three possible orders of the problem-and-solution texts, such that there were three versions of testing packets. The other comparison passages appeared in the same order: patriots vs. loyalist women, continental vs. British army, and then Mount Rushmore vs. Easter Island.

**Scoring**

**Total Main Idea Score.** For each text, students received a main idea score. Main idea scores were the total number of high level ideas recalled in the main idea, and were based on the content structure. Main ideas are those ideas which are at the top of the macrostructure. These scores vary as function of the number of ideas present in the text. The equivalent texts each had a total of 43 main ideas. For the other passages, the patriot
vs. loyalist women text had 27 ideas, the continental vs. British army text had 49 ideas, and the Mount Rushmore vs. Easter Island had 113 ideas. Each participant received a total main idea score, which was the total number of main level ideas recalled summed across all texts. For each text for which no main idea was present, students received a score of 0 ideas recalled. The scoring from the original pilot study was used. Ten percent of the sample was randomly selected, and these students’ main ideas were also scored by the author. The percentage of rater agreement was 0.99.

**Top-level-structure (TLS) score.** Top level structure scores were determined using the same procedures as Experiment 1. The scoring scale used was the same, with scores ranging from 1 to 9. The scoring descriptions used in Experiment 2 reflect the comparison structure and are derived from Meyer and colleagues’ (Meyer et al. 1980; Meyer, 1985; Meyer et al. 2002; Meyer et al. 2010; Meyer, Wijekumar, and Lin, 2001) top-level-structure scale. Scores 6-9 indicate an awareness of the comparison text structure, with a structure 9 indicating use of comparison with signaling (see Table 3-3). Students received a TLS score for each main idea, and then a mean TLS score was calculated for each student. For each text for which no main idea was present, students received a TLS score of 0. All main ideas were scored by the author. Ten percent of the sample was randomly selected and these students’ main ideas were scored by a second rater. The percentage of rater agreement was 0.86.
### Table 3-4. Comparison Main Idea TLS Scale (adapted from Meyer et al., 2011)

<table>
<thead>
<tr>
<th>Score</th>
<th>Description of TLS Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No correspondence (“I don’t know,” “I don’t remember”) or 90% or more of ideas are not from the article.</td>
</tr>
<tr>
<td>2</td>
<td>Descriptive list of ideas with no comparison of differences between two topics or creatures.</td>
</tr>
<tr>
<td>3</td>
<td>Focused only on similarities with no comparison of differences. Or, main idea is more than a descriptive list (uses another structure, such as cause-and-effect) with no ideas organized as a comparison.</td>
</tr>
<tr>
<td>4</td>
<td>Overall organization of the main idea is a list or main idea only states two topics are different, but does not contain an issue(s) on which they differ. Main idea does not compare differences of two topics on a parallel issue. Main idea main is organized like a list but a difference between topics is indicated in one part of the list.</td>
</tr>
<tr>
<td>5</td>
<td>Organized using other structure, but there is evidence that they know that the two topics were compared on at least one issue.</td>
</tr>
<tr>
<td>6</td>
<td>Organized as a comparison without signaling words (signaling words must not provide novel information, e.g., larger when comparative details are not provided); at least one parallel issue(s) of the two topics must be provided.</td>
</tr>
<tr>
<td>7</td>
<td>Organized as a comparison with one or more signaling words signaling the first comparison (e.g. different or compared).</td>
</tr>
<tr>
<td>8</td>
<td>Organized as a comparison, with a description of the first topic followed by a comparison signaling word (e.g. while,” “in contrast,,” “on the other hand”) which serves as transition to the description of the other topic.</td>
</tr>
<tr>
<td>9</td>
<td>Organized as a comparison of different topics which meets the criteria for a TLS of 7 and a TLS of 8. Or, organized as comparison with a list of features/attributes/issues on which the two topics are compared.</td>
</tr>
</tbody>
</table>

**Competency in structure scores.** Competency scores were determined in the same method as Experiment 1. The rating scales were the same with scores ranging from
1 (no use of the structure), to 6 (use of the structure); however the descriptors were different, reflecting the comparison structure. The competency rating scale is derived from Meyer et al.’s (2010) competency scale. Competency scores of 5 and 6 indicate correct identification of the entities being compared and at least one issue on which they are compared (see Table 3-4) Students received a competency score for each main idea, and then a mean competency score was calculated for each student. For each text for which no main idea was present, students received a competency score of 0. All main ideas were scored by the author. Ten percent of the sample was randomly selected and these students’ main ideas were scored by a second rater. The percentage of rater agreement was 0.89.
### Table 3-5. Comparison Main Idea Competency Scale (from Meyer et al., 2010)

<table>
<thead>
<tr>
<th>Quality</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Main idea contains no comparison of the two general or specific ideas/types from text; or, it may compare but never mentions what ideas are being compared (e.g. uses pronoun “they”).</td>
</tr>
<tr>
<td>2</td>
<td>Main idea contains an incorrect comparison of two ideas/types or a tangential text comparison that’s not the main idea.</td>
</tr>
<tr>
<td>3</td>
<td>Main idea is too general; it compares the two ideas but does not contain the specific topic/types mentioned in the text.</td>
</tr>
<tr>
<td>4</td>
<td>Main idea compares two correct, specific topics/types, but no accurate issues are compared.</td>
</tr>
<tr>
<td>5</td>
<td>Main idea compares two correct, specific topics/types on at least one accurate issue</td>
</tr>
<tr>
<td>6</td>
<td>Main idea compares two correct, specific topics/types on at least 2 issues, and one of these issues is a superordinate issue.</td>
</tr>
</tbody>
</table>

### Missing Data

The sample of participants contained a large amount of missing data, with 34% (n = 150) of students having at least one missing value. The number of students with at least one incomplete task was consistent across grade groups, with grades 4, 7, 9 having approximately 40% of students with missing data. The percentage of sixth grade students with missing values was lower (23%). In addition, it did not appear that the data was missing at random. Most of the students with missing data were in the low and middle reading groups, (n = 117, 78%), indicating that less skilled readers were more likely to
have missing data, than skilled readers. This was confirmed by an independent t-test on raw GSRT scores which showed that students with missing data had lower reading performance in comparison to students with complete data sets ($t(453) = 3.815, p < .0005$).

Further analysis was conducted in order to determine whether patterns in missing data existed and whether data analysis could minimize the amount of missing values included. Given the number of experimental tasks that students were asked to do within one sitting, there was concern that students may have experienced fatigue and/or frustration which may have contributed to non-completion of items. An analysis of missing values was conducted in order to examine at which point to stop scoring in order to minimize the amount of missing values and obtain scores representative of students’ abilities. For the problem-and-solution recalls, the largest number of missing values occurred on the final item (task 8 in the packet). For the comparison main ideas, the last two comparison items (task 7 and 8) had much larger frequencies of missing data in comparison to the other tasks. Figure 3-1 contains the frequencies of missing values across experimental tasks. These points of increased frequencies of missing values were used to establish a cut point after which no tasks would be scored. For each student, tasks one through six were scored.
Figure 3-1. Incomplete Items for Experimental Tasks

In order to provide additional support for this cut point, paired t-tests were conducted to compare mean TLS scores with and without the inclusion of these tasks for each experiment. This analysis was conducted on the original scoring of the pilot data. For Experiment 1, there was no significant difference between mean performance on the first two problem-and-solution recalls and performance on all three recall tasks ($t(437) = 0.65, p = 0.51$). This indicated that the last recall could be eliminated from analysis without losing information about students' performance. For Experiment 2, there was a significant difference in mean TLS performance on the first four main ideas in comparison to TLS scores on all comparison tasks ($t(437) = 5.71, p < .0005$), indicating that mean TLS on the first four tasks was higher in comparison to the overall mean. This result suggested that a large amount of missing data in the last two comparison tasks could be depressing TLS scores, and that the elimination of these tasks from analysis could provide a better indicator of students’ performance.
As a result of this analysis of missing data, the total number of ideas recalled, mean TLS, and mean competency scores for Experiment 1 were calculated using participants’ first and second recall. For Experiment 2, participants’ first four main ideas were used to calculate the total number of main ideas recalled, mean TLS, and mean competency scores. Because some of the comparison texts were not counterbalanced, the Mount Rushmore vs. Easter Island text was excluded from analysis.
Chapter 4

Results

Experiment 1

A multivariate analysis of variance (MANOVA) was conducted on the total number ideas recalled, mean top-level structure (TLS) scores, and mean competency scores for the problem-and-solution recalls. Predictor variables were grade (4, 6, 7, and 9) and reading skill group (low, middle, and high). The results of the multivariate analysis indicated significant main effects for grade, (Wilks' Λ= 0.84), F(9, 1030) = 8.00, \( p < .0005 \), and reading group, (Wilks' Λ = 0.78), F(6, 846) = 18.60, \( p < .0005 \). The grade by reading group interaction was not significant, (Wilks' Λ = 0.940), F(18, 1196.91) = 18.60, \( p = 0.96 \). Given the presence of significant multivariate main effects, univariate analyses (ANOVAs) were conducted on total number of ideas recalled, TLS scores, and competency scores. The alpha level was set at .05 for all statistical tests. Exact p-values are reported, except when \( p < .005 \).

Text Comprehension

The total number of ideas recalled was used as an indicator of text comprehension. A two factor analysis of variance (ANOVA) conducted on the total number of ideas recalled revealed a significant main effect for grade, F(3, 425) = 14.95, \( p < .0005 \), and reading group, F(2, 425) = 35.26, \( p < .0005 \). Post-hoc comparisons of reading groups and grades were conducted using Tukey HSD. An analysis of reading groups found that reading skill was associated with the number of ideas recalled. The high reading group \( (M = 41.90, SD = 1.51) \) recalled significantly more ideas, than middle \( (M = 33.64, SD = \)
1.27) and low ($M = 23.42, SD = 1.60$) reading groups ($p < .0005$), and the middle reading group recalled more than the low group, $p < .0005$. See Table 4-1 for reading group means for each grade. There were also significant differences among the grades in the number of ideas recalled. Sixth grade readers ($M = 35.15, SD = 1.30$) recalled more ideas than fourth grade readers ($M = 24.35, SD = 1.51$), $p < .0005$. Seventh grade readers ($M = 33.35, SD = 1.93$) recalled more ideas than fourth graders, $p = .001$. Similarly, ninth grade readers ($M = 39.09, SD = 1.95$) recalled more ideas in comparison to fourth graders, $p < .0005$. Ninth grade readers also recalled more ideas than seventh graders $p = .021$. There were no other significant differences between grades.
Table 4-1. Total Ideas Recalled, TLS, and Competency for Problem Solution Texts

<table>
<thead>
<tr>
<th>Grade 4</th>
<th>Total Ideas</th>
<th>TLS</th>
<th>Competency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M  SD</td>
<td>M  SD</td>
<td>M  SD</td>
</tr>
<tr>
<td>1</td>
<td>15.55 2.85</td>
<td>2.09 0.29</td>
<td>1.82 0.22</td>
</tr>
<tr>
<td>2</td>
<td>22.91 2.23</td>
<td>2.79 0.23</td>
<td>2.53 0.17</td>
</tr>
<tr>
<td>3</td>
<td>34.61 2.73</td>
<td>3.68 0.28</td>
<td>3.47 0.21</td>
</tr>
<tr>
<td>M</td>
<td>24.36 1.51</td>
<td>2.85 0.16</td>
<td>2.61 0.12</td>
</tr>
</tbody>
</table>

Grade 6

<table>
<thead>
<tr>
<th></th>
<th>M  SD</th>
<th>M  SD</th>
<th>M  SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25.14 2.27</td>
<td>2.84 0.23</td>
<td>2.41 0.18</td>
</tr>
<tr>
<td>2</td>
<td>34.60 2.32</td>
<td>4.11 0.24</td>
<td>3.79 0.18</td>
</tr>
<tr>
<td>3</td>
<td>45.72 2.15</td>
<td>5.26 0.22</td>
<td>4.41 0.17</td>
</tr>
<tr>
<td>M</td>
<td>35.15 1.30</td>
<td>4.07 0.13</td>
<td>3.54 0.10</td>
</tr>
</tbody>
</table>

Grade 7

<table>
<thead>
<tr>
<th></th>
<th>M  SD</th>
<th>M  SD</th>
<th>M  SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23.67 3.58</td>
<td>2.64 0.37</td>
<td>2.48 0.28</td>
</tr>
<tr>
<td>2</td>
<td>31.55 3.04</td>
<td>4.07 0.31</td>
<td>3.33 0.24</td>
</tr>
<tr>
<td>3</td>
<td>44.83 3.42</td>
<td>5.39 0.35</td>
<td>4.63 0.27</td>
</tr>
<tr>
<td>M</td>
<td>33.35 1.94</td>
<td>4.03 0.20</td>
<td>3.48 0.15</td>
</tr>
</tbody>
</table>

Grade 9

<table>
<thead>
<tr>
<th></th>
<th>M  SD</th>
<th>M  SD</th>
<th>M  SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29.33 3.86</td>
<td>3.97 0.40</td>
<td>2.94 0.30</td>
</tr>
<tr>
<td>2</td>
<td>45.50 2.53</td>
<td>5.25 0.26</td>
<td>4.35 0.20</td>
</tr>
<tr>
<td>3</td>
<td>42.43 3.58</td>
<td>5.26 0.37</td>
<td>4.55 0.28</td>
</tr>
<tr>
<td>M</td>
<td>39.09 1.95</td>
<td>4.83 0.20</td>
<td>3.95 0.15</td>
</tr>
</tbody>
</table>

Note. 1= low group, 2 = middle group, 3= high group. M = Average of all participants at each grade level collapsed over reading skill groups.

Awareness of Text Structure

Readers’ mean TLS and competency scores were used as measures of structural awareness. Two separate ANOVAs were conducted, one on TLS scores and another on competency scores, in order to determine the effect of grade and reading group on awareness of text structure. See Table 4-1 for mean TLS and competency scores for reading groups within each grade. An ANOVA of mean TLS indicated a main effect for grade, F(3, 425) = 22.81, p < .0005, and a main effect for reading group, F(2, 425) =
39.62, \( p < .0005 \). Post-hoc comparisons of reading groups and grades were conducted using Tukey HSD. An examination of differences between reading groups showed that the high reading group (\( M = 4.90, SD = 0.16 \)) had greater TLS scores than both the middle (\( M = 4.05, SD = 0.13 \)) and low groups (\( M = 2.89, SD = 0.16 \)) (\( p < .0005 \)), and the middle group performed significantly better than the low group, \( p < .0005 \).

Analysis of grade performance demonstrated that readers in later grades showed better performance in comparison to fourth graders. Sixth graders (\( M = 4.07, SD = 0.13 \)) had significantly higher TLS scores than fourth graders (\( M = 2.85, SD = 0.16 \)), \( p < .0005 \). Similarly, seventh grade (\( M = 4.03, SD = 0.20 \)) and ninth grade (\( M = 4.83, SD = 0.20 \)) readers had significantly higher TLS scores in comparison to fourth graders, \( p < .0005 \). An examination of differences among grades 6, 7, and 9 showed significant differences between ninth and sixth grade (\( p = .001 \)) and ninth and seventh grades (\( p = .006 \)), with no significant difference between seventh and sixth grade readers on TLS scores, \( p = 1.0 \). Figure 4-1 illustrates the mean grade performance on TLS.
Figure 4-1. Top Level Structure (TLS) Scores by Grade for Problem Solution Texts

An analysis of competency scores provided similar results. There was a significant main effect for grade $F(3, 425) = 19.99, p < .0005$ and reading group, $F(2, 425) = 59.14, p < .0005$. Figure 4-2 contains reading group mean competency scores for each grade. As with TLS, post-hoc comparisons of reading and grade group means were conducted using Tukey HSD. An analysis of reading group performance indicated that the high ($M = 4.27, SD = 0.12$) and middle ($M = 3.50, SD = 0.10$) groups had higher competency scores than the low group ($M = 2.41, SD = 0.12$) ($p < .0005$), and that the high group performed better than the middle group, $p < .0005$. An analysis of grade differences on competency scores provided results similar to the examination of TLS scores. Readers in sixth grade ($M = 3.54, SD = 0.10$) had higher competency scores than readers in fourth grade ($M = 2.61, SD = 0.18$), $p < .0005$. Similarly, seventh grade readers
(M = 3.48, SD = 0.15) had higher competency scores than fourth graders, p < .0005. Ninth grade readers (M = 3.95, SD = 0.15) also had higher competency scores in comparison to fourth graders, p < .0005. In addition, ninth graders performed significantly better than sixth (p = .016) and seventh graders (p = .021), while the there was no significant difference between seventh and sixth grade on competency using the problem solution structure, p = .975.

Figure 4-2. Competency Scores by Reading Group and Grade for Problem Solution Texts

Experiment 2

The analyses conducted for Experiment 2 were the same as those conducted for Experiment 1. A multivariate analysis of variance (MANOVA) was conducted on the
total number main ideas recalled, mean top level structure scores, and mean competency scores for comparison main ideas. Like Experiment 1 grade (4, 6, 7, and 9) and reading group (low, middle, and high) served as between subjects factors. The results of the multivariate analysis indicated significant main effects for grade, (Wilks' Λ = .852), F(6, 846) = 11.71, \( p < .0005 \), and reading group, (Wilks' Λ = .907), F(9, 1029) = 4.70, \( p < .0005 \). The grade by reading group interaction was not significant, (Wilks' Λ = .953), F(18,1197) = 1.13, \( p = .312 \). Given the presence of significant multivariate main effects, the univariate ANOVAs were examined for ideas recalled, TLS scores, and competency scores.

**Comprehension of Text**

The total number of main ideas recalled across the four texts analyzed was used an indicator of text comprehension. A two factor analysis of variance (ANOVA) conducted on the total number of main ideas recalled demonstrated a main effect for grade, F(3, 425) = 4.64, \( p = .003 \), and reading group, F(2, 425) = 13.24, \( p < .0005 \). Post-hoc comparisons of reading groups and grades were conducted using Tukey HSD. An analysis of reading groups found that the high reading group (\( M = 20.22, SD = 0.96 \)) recalled significantly more ideas, than middle (\( M = 15.76, SD = 0.81 \)) and low (\( M = 13.15, SD = 1.02 \)) reading groups (\( p < .0005 \)), but there was no significant difference between the middle and low groups, \( p = .260 \). See Table 4-2 for reading group means for each grade. An analysis of grade performance, yielded few significant differences between the grade groups. The sixth graders (\( M = 18.80, SD = 0.83 \)) recalled more main ideas than fourth graders (\( M = 14.88, SD = 0.96 \), \( p = .009 \). Seventh grade readers (\( M = 17.22, SD = 1.02 \))...
14.36, $SD = 1.24$) recalled fewer ideas than sixth graders ($p = .011$). There were no other significant differences between groups (see Table 4-2 for grade group means).

Table 4-2. Main Ideas Recalled, TLS, and Competency for Comparison Texts

<table>
<thead>
<tr>
<th></th>
<th>Main Ideas</th>
<th></th>
<th>TLS</th>
<th></th>
<th>Competency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Grade 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>12.09</td>
<td>1.82</td>
<td>2.11</td>
<td>0.19</td>
<td>1.25</td>
</tr>
<tr>
<td>2</td>
<td>14.78</td>
<td>1.42</td>
<td>2.60</td>
<td>0.15</td>
<td>1.52</td>
</tr>
<tr>
<td>3</td>
<td>17.78</td>
<td>1.74</td>
<td>2.94</td>
<td>0.18</td>
<td>2.00</td>
</tr>
<tr>
<td>4</td>
<td>14.88</td>
<td>0.97</td>
<td>2.55</td>
<td>0.10</td>
<td>1.59</td>
</tr>
<tr>
<td>Grade 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>17.42</td>
<td>1.45</td>
<td>2.56</td>
<td>0.15</td>
<td>1.37</td>
</tr>
<tr>
<td>2</td>
<td>18.02</td>
<td>1.48</td>
<td>3.07</td>
<td>0.15</td>
<td>1.94</td>
</tr>
<tr>
<td>3</td>
<td>20.97</td>
<td>1.37</td>
<td>3.48</td>
<td>0.14</td>
<td>2.35</td>
</tr>
<tr>
<td>4</td>
<td>18.80</td>
<td>0.83</td>
<td>3.04</td>
<td>0.09</td>
<td>1.89</td>
</tr>
<tr>
<td>Grade 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10.57</td>
<td>2.28</td>
<td>2.33</td>
<td>0.24</td>
<td>1.36</td>
</tr>
<tr>
<td>2</td>
<td>12.31</td>
<td>1.94</td>
<td>2.53</td>
<td>0.20</td>
<td>1.60</td>
</tr>
<tr>
<td>3</td>
<td>20.22</td>
<td>2.18</td>
<td>3.28</td>
<td>0.22</td>
<td>2.16</td>
</tr>
<tr>
<td>4</td>
<td>14.37</td>
<td>1.24</td>
<td>2.72</td>
<td>0.13</td>
<td>1.71</td>
</tr>
<tr>
<td>Grade 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>12.50</td>
<td>2.47</td>
<td>2.28</td>
<td>0.25</td>
<td>1.40</td>
</tr>
<tr>
<td>2</td>
<td>17.95</td>
<td>1.61</td>
<td>3.28</td>
<td>0.17</td>
<td>2.38</td>
</tr>
<tr>
<td>3</td>
<td>21.95</td>
<td>2.28</td>
<td>3.69</td>
<td>0.24</td>
<td>2.68</td>
</tr>
<tr>
<td>4</td>
<td>17.47</td>
<td>1.24</td>
<td>3.08</td>
<td>0.13</td>
<td>2.15</td>
</tr>
</tbody>
</table>

Note. 1 = low group, 2 = middle group, 3 = high group: $M =$ Average of all participants at each grade level collapsed over reading skill groups.

**Structural Awareness**

As in Experiment 1, mean TLS and competency scores served as indicators of readers’ structural awareness. Two separate ANOVAs were conducted, one on TLS scores and another on competency scores, in order to determine the effect of grade and reading group on awareness of the comparison text structure. An ANOVA of mean TLS
scores demonstrated a main effect for grade, $F(3, 425) = 4.64$, $p = .003$, and a main effect for reading group, $F(2, 425) = 13.24$, $p < .0005$. Post-hoc comparisons of reading groups and grades groups were conducted using Tukey HSD. An analysis of reading group means found differences among all three reading groups. The high reading group ($M = 3.35$, $SD = 0.10$) had higher TLS scores than the middle ($M = 2.87$, $SD = 0.08$) and low groups ($M = 2.32$, $SD = 0.10$), $p = .001$, and the middle group performed significantly better than the low group $p < .0005$. Table 4-2 contains reading group means in each grade. An analysis of grade groups revealed some differences between the grades. The sixth graders ($M = 3.04$, $SD = 0.09$) and ninth graders ($M = 3.08$, $SD = 0.13$) had higher TLS scores than the fourth graders ($M = 2.55$, $SD = 0.10$), $p = .001$. The ninth graders also performed better than the seventh graders ($M = 2.72$, $SD = 0.13$), $p = .047$. There were no other significant differences between grade groups. See Figure 4-3 for mean TLS scores for each grade.
Figure 4-3. Top Level Structure (TLS) Scores by Grade for Comparison Texts

Similar to TLS, an ANOVA on mean competency scores indicated a main effect for grade $F(3, 425) = 7.32, p < .0005$ and reading group, $F(2, 425) = 35.53, p < .0005$. Figure 4-4 contains reading group mean competency scores for each grade. An analysis of reading group performance indicated differences in structural awareness among reading groups. The high ($M = 2.30, SD = 0.08$) and middle ($M = 1.86, SD = 0.07$) reading groups had higher competency scores than the low group ($M = 1.34, SD = 0.08$) ($p = .001$), and the high group performed better than the middle group, $p < .0005$. An analysis of grade differences on competency scores found grade differences similar to those found on TLS scores. In comparison to fourth graders ($M = 1.59, SD = 0.08$), sixth graders ($M = 1.89, SD = 0.07$) and ninth graders ($M = 2.15, SD = 0.10$) had higher competency scores, $p = .011$ and $p < .0005$ respectively. Unlike grade performance on
TLS scores, ninth graders had significantly higher scores in comparison to seventh graders ($M = 1.71, SD = 0.10$) ($p = .001$) and sixth graders ($p = .019$). There were no other significant differences in competency scores between grades.

Figure 4-4. Competency Scores by Reading Group and Grade for Comparison Texts
Chapter 5
Discussion of Results

The results of both experiments indicate that structural awareness and text comprehension increase with both the grade and the overall reading comprehension ability of the learner. In general, more skilled readers recalled more ideas and demonstrated greater performance on structural awareness measures in comparison to less skilled readers. Similarly, older readers demonstrated greater recall and structural awareness than younger readers. These findings were relatively consistent across both experiments, suggesting that problem-and-solution and comparison text structures follow similar developmental patterns. In regards to the three hypotheses tested, the lack of a significant skill by grade interaction provides evidence in favor of the stable difference hypothesis over the increased difference hypothesis. There does not appear to be a Mathew effect in structural awareness; differences in reading ability groups were consistent across grades. In addition, the presence of significant grade differences between the earlier grades (grades 6 and 4) on all measures provides support for the stable difference hypothesis over the no difference hypothesis as well. Of the three hypotheses posed, the results of the study support the stable difference hypothesis as the best explanation of the relationship between age and skill. However, further examination of grade means suggests that the relationship between age and structural awareness is complex, and that the stable difference hypothesis may not adequately account for the development of structural awareness.

Comprehension of Text
For the problem-and-solution texts used in Experiment 1, the sixth, seventh, and ninth grade students all recalled more ideas from the texts in comparison to the fourth grade students. However, there were few significant differences within this late elementary to high school group. There was no significant difference in recall between sixth and seventh grade, and while ninth graders recalled more than sixth and seventh graders, only the difference between the ninth and seventh was significant. These results suggest that although the older (late elementary through high school) readers had better comprehension in comparison to earlier elementary school readers, across these later grades there was very little difference in text comprehension.

Unlike Experiment 1, the results of Experiment 2 indicated very little difference in the ability to write a main idea for a comparison text across all of the grades. Sixth grade readers included more main ideas than fourth graders and seventh graders, but there were no other significant differences between any of the other grades. Although these results seem to support a no difference hypothesis, no differences in main ideas scores with age could be due to the nature of the task. Unlike the recall task of Experiment 1, the main idea task of Experiment 2 provided students with limited space and the text in view while composing a two-sentence main idea. Moreover, the task asked them to write only those ideas central to the text. As a result, the number of ideas students would be likely to write would be smaller, probably making the measure less sensitive to grade group differences.

**Structural Awareness**

Similar to text comprehension, grade group performance on measures of structural awareness indicated an increase in structural awareness with age. However,
there were instances at which no increase with age occurred. In Experiment 1, all of the older readers (grades 6-9) demonstrated greater structural awareness for the problem-and-solution structure in comparison to grade 4 readers. Similar to recall, there was no difference with between sixth and seventh graders. Unlike recall, ninth graders had significantly higher TLS scores than both the sixth and seventh graders. Competency ratings provided similar results. Both measures indicate an increase in structural awareness at sixth grade, no difference between sixth and seventh, and then an increase in awareness at ninth in comparison to both sixth and seventh grade.

Performance on the structural awareness measures of Experiment 2 provided somewhat similar results; however, the pattern of grade differences was more complex. For the TLS scores, only sixth and nine graders demonstrated significantly greater awareness of the comparison structure in relation to fourth graders. In contrast to Experiment 1, the only significant difference in TLS among the sixth, seventh, and ninth grade groups was the increase in ninth grade performance in comparison to seventh grade performance. Unlike the pattern found for TLS of main ideas, grade differences in competency scores in Experiment 2 reflected a very similar pattern to that found in Experiment 1. Sixth and ninth graders had significantly greater levels of structural awareness than fourth graders, and ninth grade readers had greater structural awareness in comparison to both sixth and seventh grade readers. Similar to Experiment 1, the difference between sixth and seventh grade performance was not significant.

The results of both Experiment 1 and 2 suggest that although structural awareness does develop with age, this development may not be linear. In the late elementary through middle school years there may be few gains in comprehension and structural
awareness. As a result, the stable difference hypothesis may fail to provide a sufficient explanation of the development of structural awareness. In this particular sample, there appears to be little differences in both text comprehension and structural awareness between sixth and seventh grades. This lack of a difference between sixth and seventh grade, may be related to little change in overall comprehension. In this particular sample the sixth and seventh grades performed similarly on the GSRT. Overall, the results of the study suggest that increases in structural awareness will occur to the extent that changes in overall reading ability are present. When there is little change in overall reading ability, there is little change in structural awareness.

**Limitations of the Study**

Despite their implications for the nature of both structural awareness and comprehension development, the results of this study should be interpreted cautiously, particularly any conclusions regarding growth. Because of both the cross-sectional design and the small number seventh and ninth grade students included in the study, it may be possible that in some instances there was not enough statistical power to detect differences between groups. Additionally, given that the majority of students excluded from the study were from the seventh grade, it is possible that problems in the administration of the study could have contributed to the results. Also, given that sixth and seventh grade students had similar overall reading comprehension performance, it may be possible that the seventh grade sample used in this study is not representative of typical seventh grade reading abilities.

**Implications for Future Research**
Additional research is needed in order to determine whether the patterns found with the sixth, seventh, and ninth grade participants of this study would replicate with different readers and different aspects of reading comprehension. The results of this study indicate that as students move from late elementary school through to high school, both reading comprehension and structural awareness development may slow down, with little change occur during this time. It is possible that the results of this study reflect the students in this particular school district and are not indicative of general comprehension development. Additional research with young adolescents, particularly longitudinal research, is needed in order to determine how comprehension develops throughout this time and how this development may differ from that of the early elementary school years. The findings of this study do indicate that development in these later years is complex, and a hypothesis which rest on assumptions of linear growth in comprehension may be inappropriate; alternative approaches are needed.

It is also important to note that while structural awareness increased with age, no reading group in any grade had mean TLS or competency scores that would indicate that readers were structurally aware. Although readers may develop structural awareness, in order to achieve mastery in their abilities to recognize and use structure, explicit instruction is likely needed. The results from this study indicate that readers in all of the grades could have benefited from text structure interventions, even those readers classified as highly skilled readers. Moreover, given the finding that there was no difference between sixth and seventh grader readers and neither group demonstrated competency in structural awareness, this study suggests that middle school students may be in particular need of explicit instruction in the use of text structure.
References


Appendix A

Problem and Solution Texts

Cat Text (Copyright © 2004, B.J.F. Meyer)

Felines often become fat, especially cats that live their lifetimes in homes with cat fanciers and families. This is a real hazard for a cat that spends a lifetime confined inside of only one home. Obesity can be a reaction to boredom, a bored brain, in an indoor cat.

At a meeting sponsored by the American Veterinary Council, Dr. Ann G. N. Franz, an Australian veterinarian, recommended fresh air and bird viewing via an all season cat stroller. Cat fanciers who provide bird watching and routine strolling for aging cats often stimulate their cats' brains and healthier eating.

Dog Text (Copyright © 2004, B.J.F. Meyer)

Canines who like the smell and taste of cocoa bean shells often become sick eating them. This is a real hazard when dogs run into gardens where mulch is made from chocolate-making leftovers. Poisoning is a reaction in dogs to theobromine in cocoa bean shells.

At a meeting sponsored by Animal Poison Control, Dr. R. F. S. Drolet, a Canadian pathologist, recommended avoiding cocoa bean mulch and immediate treatment of poisoning symptoms by dog owners. Owners who keep cocoa bean shells out of their yards and firmly control leashed dogs can effectively prevent tremors and threat of death.
Psychologists who work with rats and mice in experiments often become allergic to these creatures. This is a real hazard for these investigators who must spend hours a week running rats in experiments. These allergies are a reaction to the protein in the urine of these small animals.

At a meeting sponsored by the National Institutes of Health, Dr. Andrew J. M. Slovak, a British physician, recommended kindness to rats and mice by the experimenters. Psychologists who pet and talk softly to their rats are less often splattered with urine and the protein that causes the allergic reaction.
Appendix B

Problem and Solution Content Structures

(Copyright © 2004, B.J.F. Meyer)

Cat Text Structure

response

problem REAL HAZARD
BECOME FAT
manner
OFTEN
patient
FELINES
description: equivalent
CATS
description: attribution
LIVE WITH
patient
collection
CAT FANCIERS
FAMILIES
range
HOMES
description: specific
CONFINED INSIDE ONLY ONE HOME

setting: time

(THEIR) LIFETIME(S)

causation: explanation CAUSES

CAN BE A REACTION

patient, latter

OBESITY specific word

former

BOREDOM

description: specific

BORED BRAIN emphasizes brain

range

INDOOR CAT (CAT NOT ALLOWED OUTSIDE)

solution recommended suggested

RECOMMENDED

patient

collection

FRESH AIR

BIRD VIEWING

instrument

ALL SEASON CAT STROLLER

agent (they say, but not passage says)

DR. FRANZ
description: specific

ANN G. N.

description, attribution

collection

AUSTRALIAN VETERINARIAN

description: setting

SPONSORED

patient

MEETING

agent

AMERICAN VETERINARY COUNCIL

causation: explanation

causation: covariance: antecedent

collection

PROVIDE BIRD WATCHING

PROVIDE STROLLING

manner

ROUTINE

agent

CAT FANCIERS

patient

AGING CATS
causation: covariance: consequent

STIMULATE

manner

OFTEN

range

collection

CATS' BRAINS

HEALTHIER EATING

Dog Text Structure

response

problem REAL HAZARD

BECOME SICK EATING COCOA BEAN SHELLS

manner

OFTEN

patient

CANINES

description: equivalent

DOGS

description: attribution

LIKE

patient
collection
SMELL
TASTE
range
COCOA BEAN SHELLS
description: specific

MULCH MADE OF CHOCOLATE MAKING LEFTOVERS
setting: time
RUN IN TO GARDENS
causation: explanation CAUSES

IS A REACTION
patient, latter
POISONING
former
EATING COCOA BEAN SHELLS
description: specific
THEOBROMINE
range
DOGS
solution recommended suggested
RECOMMENDED
patient
collection
AVOIDING COCOA BEAN (SHELL) MULCH

IMMEDIATE TREATMENT OF POISONING SYMPTOMS

instrument

DOG OWNERS

agent

DR. DROLET

description: specific

ROBERT F. S.

description, attribution

collection

CANADIAN

PATHOLOGIST

description: setting

SPONSORED

patient

MEETING

agent

ANIMAL POISON CONTROL CENTER

causation: explanation

causation: covariance: antecedent

collection

KEEP COCOA BEAN SHELLS OUT OF THE YARD

CONTROL
manner
FIRMLY
agent
OWNERS
patient
LEASHED DOGS
causation: covariance: consequent
CAN PREVENT
manner
EFFECTIVELY
range
collection
TREMORS FROM THEOBROMINE
THREAT OF DEATH
Rat Text Structure
response
problem REAL HAZARD
BECOME ALLERGIC TO CREATURES
manner
OF TEN
patient
PSYCHOLOGISTS
description: equivalent

INVESTIGATORS

description: attribution

WORK WITH

patient

collection

RATS score

MICE

range

EXPERIMENTS

description: specific

RUNNING RATS IN EXPERIMENTS

setting: time

HOURS A WEEK

causation: explanation

ARE A REACTION

patient, latter

ALLERGIES

former

URINE

description: specific

PROTEIN

range
THESE SMALL ANIMALS

solution recommended suggested

RECOMMENDED KINDESS TO

patient
collection

RATS
MICE

instrument

EXPERIMENTERS/PSYCHOLOGISTS/INVESTIGATORS

agent

Dr. SLOVAK
description: specific

ANDREW J. M.
description; attribution
collection

BRITISH

PHYSICIAN doctor
description: setting

SPONSORED

patient

MEETING

agent

NATIONAL INSTITUTES OF HEALTH
causation: explanation

causation: covariance: antecedent

collection

PET

TALK TO

manner

SOFTLY

agent

PSYCHOLOGISTS

patient

RATS mice

causation: covariance: consequent

ARE LESS SPLATTERED WITH

manner

OFTEN

range

collection

URINE

PROTEIN THAT CAUSES ALLERGIC REACTION
Appendix C

Comparison Texts

Penguin Text (Copyright © 2004, B.J.F. Meyer)

Emperor penguins and Adelie penguins are ___________ from one another. Emperor penguins are the largest of all the penguins. They may grow up to 4 feet tall and can weigh more than 90 pounds. They have yellow necks and long beaks. Emperor penguins live in icy Antarctica. The emperor penguins feed mainly on fish and squid. ___________ the large emperor penguins, Adelie penguins are ___________; some Adelie penguins grow only about 2 feet high and others reach 3 feet. An Adelie penguin has a short beak and feathers that cover most of it. They have no feathers on the tips of their beaks. Adelie penguins feed almost entirely on krill, which are part of crustaceans. The ___________ as emperor penguins, Adelie penguins live in icy Antarctica.

Primate Text (Copyright © 2004, B.J.F. Meyer)

Pygmy and tamarin primates are very ___________ from one another. Pygmy primates are the smallest of all the primates. They may grow up to 6 inches long and can weigh only 4 ounces. They are yellow brown. They have lower teeth of equal length. Pygmies live in South America. The pygmy primates feed almost entirely on gums, which are part of tree sap. ___________ the tiny pygmy primates, tamarin primates are ________; some kinds of tamarins grow to about 13 inches long and others reach 16 inches.
Tamarins have u-shaped jaws and some of their teeth are longer than others. The canine teeth are much longer than the incisors. They feed mainly on fruit and insects. The ______ pygmy primates, tamarin primates live in South America.

_Turtle Text (Copyright © 2004, B.J.F. Meyer)_

Leatherback turtles and hawksbill turtles are _______ from one another.

Leatherback turtles are the largest of all the turtles. They may grow up to 6 feet long and can weigh more than 900 pounds. They are shiny black. They have no claws. Leatherback turtles live in tropical waters. They feed almost entirely on jellyfish, which are part of ocean invertebrates.

__________ the huge leatherback turtles, hawksbill turtles are _________; some kinds of hawksbill turtles grow only about 2 1/2 feet long and others reach 3 feet. Hawksbill turtles have a hawk-like jaw and two claws on each of their front and rear flippers. The hawksbill turtle feeds mainly on sponges and squid.

The ______ as the leatherback turtles, the hawksbill turtles live in tropical waters.

_Patriot vs. Loyalist Women Text (Berson, 2003, p. 311)_

Many women in the colonies took part in the Patriots’ war effort. Some ran farms and businesses. Others formed groups to raise money for the war and to collect clothing for the soldiers. Esther Reed started the Philadelphia Association in 1780 to help the Continental Army. When Reed died, Sarah Franklin Bache (BAYCH), Benjamin Franklin’s daughter, took over the Association.
Not all women were Patriots, __________. There were Loyalist women in every colony. Some fought for the British. Many others gave the British food and other supplies

**Continental vs. British Army Text (Berson, 2003, p. 315)**

The Continental Army soldiers who stood before George Washington that summer day in 1775 had never fought as an army before. Some of them had fought in the French and Indian War, ________ most had no military experience. Washington quickly had to make rules for the soldiers and get them trained to fight against the British.

________ the Continental Army, the British army was made up professional soldiers. They had the best training and the most experienced officers. ________ the British had ________, too. It was difficult to fight a war more than 3,000 miles (4,828 km) from home. They often had ________ delivering soldiers and supplies across the Atlantic Ocean.
Appendix D

Comparison Content Structures

(Copyright © 2004, B.J.F. Meyer)

Penguin Text Structure

Comparison

EMPEROR PENGUINS

description: attribution

collection

ARE LARGE

range

PENGUINS

description: specific

ARE LARGEST

range

PENGUINS

description: specific

ALL

description: specific

collection

MAY GROW UP TO
latter

FEET TALL

description: specific

(FOUR) 4

CAN WEIGH

description: manner

MORE THAN

latter

POUNDS

description: specific

(NINETY) 90

HAVE NECKS

description: attribution

YELLOW

HAVE LONG BEAKS

LIVE

range

ANTARCTICA

description: specific

ICY

FEED

description: manner

MAINLY
patient
collection
FISH
SQUID

ADELIE PENGUINS

description: attribution
collection

ARE SMALL
range
PENGUINS
description: specific
comparison

GROW
agent
SOME KINDS
description: manner

ONLY
range, latter
FEET HIGH
description: specific

(TWO) 2
REACH
agent
OTHERS
range, latter

FEET
description: specific

(THREE) 3

HAVE BEAK
description: attribution

SHORT

HAVE FEATHERS
description: attribution

COVER
range

MOST OF IT
comparison

HAVE NO FEATHERS
range

TIP

FEED
description: manner

ALMOST ENTIRELY
patient

KRILL
description: constituency identification
CRUSTACEANS

LIVE

range

ANTARCTICA
description: attribution

ICY

Primate Text Structure

Comparison

PYGMY MAMMALS
description: attribution
collection

ARE TINY

range

PRIMATES
description: specific

ARE SMALLEST

range

PRIMATES
description: specific

ALL
description: specific
collection
GROW TO ABOUT
latter
INCHES LONG
description: specific
(SIX) 6
CAN WEIGH
description: manner
ONLY
latter
OUNCES
description: specific
(FOUR) 4
ARE BROWN
description: specific
YELLOW
HAVE TEETH
description: specific
LENGTH
description: attribute
EQUAL
LIVE
range
SOUTH AMERICA

FEED

description: manner

ALMOST ENTIRELY

patient

GUMS

description: constituency

TREE SAP

TAMARINS

description: attribution

collection

ARE LARGER

range

PRIMATES

description: specific

comparison

GROW

agent

SOME KINDS

description: manner

ABOUT

range, latter

INCHES LONG
description: specific

(THIRTEEN) 13

REACH
agent

OTHERS
range, latter
INCHES
description: specific

(SIXTEEN) 16

HAVE JAWS
description: specific

U-SHAPED

HAVE TEETH
description: specific
comparison

SOME LONGER
THAN OTHERS
description: explanation
comparison

CANINES LONGER
THAN INCISORS

FEED
description: manner
MAINLY

patient
collection
FRUIT
INSECTS
LIVE
range
SOUTH AMERICA

Turtle Text Structure

Comparison

LEATHERBACK TURTLES
description: attribution
collection
ARE HUGE
range
TURTLES
description: specific
ARE LARGEST
range
TURTLES
description: specific
ALL
MAY GROW UP TO
latter
FEET LONG
description: specific
six
CAN WEIGH
description: manner
MORE THAN
latter
POUNDS
description: specific
ninety
ARE BLACK
description: attribution
SHINY
HAVE NO CLAWS
LIVE
range
WATERS
description: specific
TROPICAL
FEED

description: manner

ALMOST ENTIRELY

patient

JELLYFISH

description: constituency identification

OCEAN INVERTEBRATES

HAWKSBILL TURTLES

description: attribution

collection

ARE SMALL (SMALLER)

range

TURTLES

description: specific

comparison

GROW

agent

SOME KINDS

description: manner

ONLY ABOUT

range, latter

FEET LONG

description: specific
(two and a half) 2.5 REACH
agent
OTHERS
range, latter
FEET
description: specific
(three) 3 
HAVE JAW
comparison: analogy
HAWK LIKE
HAVE CLAWS
description: specific
TWO
range
EACH FLIPPER
description: specific
collection
FRONT
REAR
FEED
description: manner
MAINLY
Patriot vs. Loyalist Women Text

Comparison HOWEVER, NOT ALL TOOK PART IN agent WOMEN description: attribution MANY PATRIOTS description: setting location COLONIES range, patient WAR EFFORT description: specific
Patriots'
description: specific
collection
ran
patient
farms
businesses
agent
some (women)
formed
patient
groups
agent
other (women)
latter
collection
raise
patient
money
latter
war
collect
patient
CLOTHING
latter

SOLDIERS
description: specific
sequence WHEN
STARTED
agent

REED
description: specific

ESTHER
patient

ASSOCIATION
description: specific

PHILADELPHIA
description: setting time
in 1780
latter

HELP
patient

ARMY
description: specific

CONTINENTAL
TOOK OVER
patient

ASSOCIATION

agent

BACHE (BAYCH)

description: specific

SARAH

description: specific

FRANKLIN

description: attribution

DAUGHTER

description: constituency identification

FRANKLIN

description: specific

BENJAMIN

WERE

range

LOYALIST

patient

WOMEN

description: setting

EVERY COLONY

description: specific

collection
FOUGHT FOR
agent
SOME (WOMEN)
range
BRITISH
GAVE
agent
OTHERS
description: attribution
MANY
patient
collection
FOOD
OTHER SUPPLIES
benefactive, latter
BRITISH

Continental vs. British Army Text Structure

Comparison UNLIKE
CONTINENTAL ARMY
description: attribution
response: problem/solution: problem
HAD NEVER FOUGHT

patient

SOLDIERS

description: specific

ARMY

description: specific

CONTINENTAL

STOOD BEFORE

range

WASHINGTON

description: specific

GEORGE

description: setting time

*1775

description: specific

THAT DAY

description: attribution

SUMMER

range

ARMY

description: setting time

BEFORE

description: specific
HAD NO
agent

MOST
patient

EXPERIENCE
description: specific

MILITARY
comparison: adversative BUT

HAD Fought
range

WAR
description: specific
collection

FRENCH
INDIAN
agent

SOME
response: problem/solution: solution
collection

HAD TO MAKE
agent

WASHINGTON
description: manner
QUICKLY

patient

RULES

benefactive

SOLDIERS

GET TRAINED

agent

WASHINGTON

patient

THEM

range

FIGHT

patient

BRITISH

BRITISH ARMY

description: attribution

WAS MADE UP

range, former

SOLDIERS

description: specific

PROFESSIONAL

description: attribution

HAD
patient
collection

TRAINING
description: attribution
BEST
OFFICERS
description: attribution
EXPERIENCED
description: attribution
THE MOST
comparison: adversative BUT
HAD PROBLEMS
description: manner
TOO
description: specific
WAS DIFFICULT TO problem, trouble
patient
FIGHT
patient
WAR
range, latter
MILES
description: specific
*3,000

description: equivalent

*4,828

description: specific

KM

former

HOME

description: specific

HAD TROUBLE

description: manner

OFTEN

patient

DELIVERING

patient

collection

SOLDIERS

SUPPLIES

range

OCEAN

description: specific

ATLANTIC