DO SECOND AND THIRD GRADE CHILDREN MAKE CAUSAL INFERENCES

WHEN READING EXPOSITORY TEXTS?

by

MATTHEW QUIRK

(Under the Direction of Paula Schwanenflugel)

ABSTRACT

Prior studies have suggested that children as young as the second grade are able to draw inferences when reading narrative texts (Casteel, 1993; Casteel & Simpson, 1991). The purpose of the present study was to investigate whether students were making inferences while reading expository texts at various age and skill levels. Participants were 64 second graders and 78 third graders from urban schools in Georgia and New Jersey. Participants were presented six-sentence expository passages. The children read pairs of sentences aloud, after which an unrelated word, an associated word, or an inference word were presented and the child read the word as quickly and accurately as possible. Results indicated that no inferences were being generated regardless of age or skill level.

INDEX WORDS: Causal Inferences, Reading Expository Text, Children Inferencing, Development of Fluency

DO SECOND AND THIRD GRADE CHILDREN MAKE CAUSAL INFERENCES

WHEN READING EXPOSITORY TEXTS

by

MATTHEW QUIRK

B.S., Pennsylvania State University, 1999

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial

Fulfillment of the Requirements for the Degree

MASTER OF ARTS

ATHENS, GEORGIA

2002

DO SECOND AND THIRD GRADE CHILDREN MAKE CAUSAL INFERENCES

WHEN READING EXPOSITORY TEXTS

by

MATTHEW QUIRK

Major Professor: Dr. Paula Schwanenflugel

Committee: Dr. Steven Stahl
Dr. Nancy Knapp

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2002

ACKNOWLEDGEMENTS

I would like to acknowledge several people who have assisted me in preparing my thesis. First, I would like to thank my parents, John and Kathy Quirk, for all of their insight and support throughout my graduate school experience. I would also like to thank my major professor Dr. Paula Schwanenflugel for all of her guidance that enabled me to make it through my Master's thesis. Finally, I would like to thank Dr. Steven Stahl and Dr. Nancy Knapp for all of their support and advice as I worked through my revisions on this project.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iv

LIST OF TABLES vi

CHAPTER

1 INTRODUCTION 1

Inferencing in Expository Text 5

Developmental Studies of Inferencing during Early Reading 9

2 METHODS 15

Design 15

Participants 15

Stimuli 15

Procedure 17

3 RESULTS 22

Naming Time Analysis 23

Accuracy Analysis 25

4 DISCUSSION 27

REFERENCES 32

APPENDICES

A APPENDIX A: TEST ITEMS AND TARGETS 34

B APPENDIX B: ITEMS ANALYSIS 37

LIST OF TABLES

Page

Table 2.1: Example Passage and its Corresponding Inference Eliciting Question............20

Table 2.2: Standard Scores from the WIAT Used to Determine Reading Skill Level 21

Table 3.1: Subject Means for Unassociated, Associated, and Inference Target Words Split by Grade and Skill Level 23

CHAPTER 1: INTRODUCTION

A key component of reading comprehension at any level is the ability to make inferences while reading. This ability allows the reader to "fill in the gaps" of what they are reading so that the text flows from idea to idea. If these gaps are not filled the text is choppy and difficult to understand.

There are many different types of inferences that a person can make while reading. Each of these different types of inferences serves a different purpose in the comprehension process. For our purposes in this study I will be looking at what is called a causal antecedent or backward inference. In order to understand what I am investigating it is essential to understand the differences between a causal (backward) inference and a predictive (forward) inference.

A causal inference essentially is one that is made in order to fill in gaps in a causal sequence that are needed to fully comprehend the ongoing text. In order to make this causal inference the reader must link prior information in the sequence to the ongoing text using their prior knowledge base in the subject area at hand. In contrast, a predictive or forward inference is one where the current information being read is used to make an inference about what will happen next or later on down the line. In this study I will be concentrating on causal rather than predictive inference making. Therefore, it is essential to clearly define causal inferences in order to gain a better understanding of what constitutes the kind of inferencing that is the focus of this study.

Millis and Graesser (1994) have broken down the pieces of a causal inference into four key components: temporal priority, operativity, necessity, and sufficiency. Temporal priority simply means that the cause must precede the consequent where the inference is necessary. The cause must also remain active at the point where the inference is needed (operativity). The necessity and sufficiency criteria are both contextual in nature. They state that the consequent could not occur without the presence of the cause prior to it (necessity) and that the presence of the cause is sufficient for the presence of the consequent (sufficiency). Causal inferences are made when these components are not explicitly stated in the story and are needed to maintain coherence in the text (Casteel, 1993).

The following passage is an example of how the CSS model applies for the stimuli in the current study:

A baby whale cannot swim very fast.

A big hungry shark is watching.

It is waiting for a chance to attack. (cause)

When his mother sees the shark, she rushes straight at it. (consequent)

Mother whales can swim fast.

The baby whale is safe again.

In this example, the shark waiting to attack is the cause and the mother rushing to her baby is the consequent. The fact that the mother is rushing to her baby to protect it from the hungry shark is the necessary inference. The cause directly precedes the consequent so it satisfies the requirement of temporal priority and operativity. The inference is necessary because the mother rushes to her baby only after seeing the shark. It is sufficient because the mother would be expected to protect her baby with a hungry shark watching.

The current-state selection model or CSS model explains how these causal connections are made between clauses (Graesser, Swamer & Baggett, 1996). This CSS model is very similar to other models that explain strategies for local text coherence (Trabasso & van den Broek, 1985).

The CSS model assumes that the reader focuses his or her attention on a small set of clauses in short term memory after each sentence is read. The reader can quickly and easily comprehend the next sentence if the causal or non-causal bridge is made between the small set of clauses in their short-term memory and the new information from the incoming sentence. If no connection can be made the reader tries to establish coherence by bridging the sentences with a causal inference. This process will require more time for the reader to bridge the sentences by making the inference. If no connection can be made then the reader may continue reading where the connection may be made further into the text. These inferences are usually used to answer a "why" question, and the causal inference that is made is usually a "because" statement that answers that question. These processes play a significant role in comprehension of the text being read.

This brings us to the role that these inferences play in the overall picture of reading comprehension. You can break comprehension into two types, macro and microcomprehension. Microcomprehension involves the ability to understand relationships within and between syntactic structures in sentences. Macrocomprehension is more affected by prior knowledge and consists of a more global perspective than the micro counterpart.

There is definitely a macro component involved because prior global knowledge plays a significant role in the ability to make causal inferences while reading expository text. If you do not have the prior knowledge necessary, how can you be expected to make the inference? For example, if I am reading a passage about the ocean and I have no prior knowledge about the ocean, my inability to make a necessary causal inference may not be due to a lack in ability. However, the cause may simply be that I do not have the knowledge necessary to make the connection. This is a problem that we must be very careful to avoid when studying inferencing in young readers. A lack in knowledge must not be mistaken for a lack of the ability to make an inference.

There is also a microcomprehension component as well that involves the relationship between sentences. In order to understand the text you are reading it is essential to understand the relationships between sentences and the syntactic structures within them. The aforementioned components of causal relationships are keys to microcomprehension. If one or more of these keys is not an explicit piece of the text it becomes increasingly necessary to make the causal inferences in order to keep the text flowing in a logical manner. Once we have established that the global knowledge necessary to make the inference is present, we will actually be measuring inference-making ability using these microstructures. In the study discussed here, the children did not read entire books; instead they read short passages from books that had missing links embedded within them where causal inferences were necessary. These passages were taken out of expository rather than narrative text.

<u>Inferencing in Expository Text</u>

Many of the studies previous to this one have looked at inference making ability in narrative rather than expository text. Also, most of them have involved adults rather than children. Although not directly bearing on the subject of the present study (the ability of young readers to draw inferences during the reading of expository text), it is definitely relevant to review what these studies have found regarding inference making abilities in adults. It has been found that when reading expository text readers make causal inferences but not predictive inferences (Millis & Graesser, 1994). The inability to make predictive inferences makes logical sense because many of the things that we read in expository text are unfamiliar. This would make it very difficult to make predictions about what will happen next if the material is unfamiliar. In narrative text this is not the case because generally the subject matter is not new; therefore, it is much easier to make a prediction about what will happen next.

In a study by Noordman, Vonk, and Kempff (1992) adult subjects read expository passages that required them to make causal inferences to maintain contextual coherence. Each subject read one of two lists containing 5 practice passages along with 16 experimental passages. The passages were split into two conditions, explicit and implicit. Both explicit and implicit passages contained a *because* sentence that was the target sentence, which required an inference to understand. In the explicit condition, the information needed to make the inference was given to the reader in the sentence directly proceeding the target sentence. This information was not provided in the implicit condition. To avoid priming effects, the explicit sentence did not contain any of the content words contained in the target sentence. The two lists contained an equal amount of passages from each condition, where the explicit form of a passage was on list one and the implicit form of passage one was on list two. Following each fourth passage was a verification task immediately following the reading (4 tasks for each subject). One of the tasks required the inference information and used content words from the explicit form of the passage, another was unrelated but also true, and the other two were false. If the inference were being made during reading, one would expect the reading times to be different for the implicit and explicit passages. If the inference were being made after reading, the reaction times for the true/false verification questions in each condition would be significantly different. It was found that under normal reading conditions, the adult readers were not making the inferences needed for coherence during reading. However, when the subjects were asked to answer questions following the reading to verify information constructed by the causal inference, they were making the inference following the reading. This would be similar to the situation that children in schools are put into almost everyday when they read required material. All questions that they are required to answer are processed following the reading and do not require the reader to make online decisions.

McKoon and Ratcliff (1992) argued that readers routinely make inferences while reading expository text, but they must be knowledge-based rather than elaborative. This view is called the minimalist hypothesis on inference generation. Support for the minimalist hypothesis was provided in three ways. First, it was shown that inferences are made for local coherence but not automatically encoded for global coherence. Basically, inferences were made to connect closely linked details in the story, but they were not made to connect more global and widely spread concepts in the story. Second, it was shown that the inferences made were necessary for understanding the text; however, elaborative inferences that added to the existing information were not made. Finally, inferences were made in life-like situations where actions followed a logical progression, but they were not made in situations where the progression of events did not follow life-like patterns. This relates back to my previous discussion on the importance of knowledge base in order to make inferences. From the minimalist perspective, it would be unlikely that children would be able to make inferences in subject matter that is unfamiliar to them. This would require the student to construct an inference from information that is either partially or completely foreign to them.

Millis and Graesser (1994) narrowed their study to only include the construction of knowledge-based inferences while reading expository text. Following each passage the subjects were given two comprehension questions to ensure that they understood what they had read. The results indicated that adults generated causal antecedent inferences but not causal consequential inferences. This was shown by comparing the two groups' average reaction times for reading either inference or unrelated words following the completion of each sentence. Interestingly, unlike Noordman et al. (1992), Millis and Graesser (1994) found that the generation of inferences was indeed happening online. The subjects were either given 540ms stimulus onset asynchrony (SOA) or 1040ms SOA before the inference or unrelated word appeared. The difference in times was used to test whether the extra time allowed for the 1040 SOA group would result in more inference generation as a result of having more time to process the inference following completion of the sentence. Because the speed to recognize inference words was similar for both the 540 SOA and 1040 SOA groups, it was determined that the inferences were being generated before the subject had finished the sentence. This would make the amount of time given to each student to process the inference irrelevant considering they have already processed the inference by the time they had completed the sentence, if inferences were being generated at all.

There are many factors to control for when designing studies assessing inferencing. It is important to understand these factors so that they do not give false indicators when interpreting the results of an instrument that has been administered. In fact, these factors should be taken into account when creating the instrument so that they are not at all involved in the results taken from that instrument.

Many of these factors and their effects are outlined in a study by Singer, Harkness, and Stewart (1997). Perhaps the most important of these factors involved the sequencing of the statements involved in the passages used to test inference generation. It has been found that ideas presented earlier in a passage are seen as more important than those toward the end of a passage. Therefore, if the causal and consequent are placed at the beginning of a passage rather than the end, there is more attention allocated to it and may cause more inferences to be made in that situation. In contrast, if the causal inference needs to be made at the end of a passage there may be less attention allocated to it and inferences may not be made at the same rate (Singer et al, 1997).

Another confounding factor that deals with the positioning of the sentences in a passage is the purpose of the first sentence. It has been found that there is evidence suggesting that reading times for the initial sentence in a paragraph are slightly higher than the times for the remaining sentences in the paragraph. Reading times are said to be slightly slower when an inference is being generated online; however, if the causal statement is the first sentence in a paragraph the reading time could be misleading. Although it may appear to be slower because an inference is being generated, it may only be slower due to the position in that passage that it is placed (first). Making sure that there is a leading sentence in each passage before the target sentences begin to appear can eliminate this problem.

Another problem that needed to be addressed in this study had to do with associated words. The inference words that we use cannot be an associate of any words in the text of the passage. If associates were used it would be impossible to determine whether the source of the reaction time difference was due to inference generation or simply because the word was associated with other words in the text making it easily accessible.

Developmental Studies of Inferencing during Early Reading

There is not much research examining the development of inferencing while reading in children who are in the process of learning to read fluently. Casteel and Simpson (1991) examined the ability of second, fifth, and eighth-grade children to make inferences when reading. The study explored both forward (predictive) and backward (causal) inferencing and also looked at whether the inferences were being generated online during encoding or offline during retrieval when needed. The children read eight stories that contained eleven sentences each. Within each story there were two forward inference sentences, two backward inference sentences, two inference inducing sentences (for the backward inferences), and five filler sentences. Following each story the subjects answered 8 yes /no question that determined if the inferences were being made. Also, their reading times for each sentence were being measured to determine if inferences were being generated online. This study found that backward inferences tended to be generated during encoding and forward inferences were generated offline during retrieval. The results also suggested that 2nd graders were making significantly fewer inferences than all other grade levels tested; however, they drew both forward and backward inferences at a rate significantly higher than chance when reading these narrative texts.

Casteel (1993) examined inferencing further by looking at the effects necessity and reading goal had on inference generation for second, fifth, and eighth grade readers. The study was conducted using four sentence narratives that were designed to elicit either a forward or backward inference. After each passage, one question was asked to determine if an inference was generated along with another filler question. Necessity was established through the use of Trabasso and van den Broek's (1985) causal chain criteria for local text cohesion. Reading goal was superficially manipulated through the directions given before the subject began the task. Again, this study found that while reading these narrative passages, children of all ages were able to make backward inferences with relative ease. This finding strongly supports their hypothesis that backward inferences were likely to be generated more frequently than forward inferences due to the fact that they are necessary to maintain local coherence of the passage. Reading goal was also found to have an effect on reading time. The integrative group was told that they would need to read carefully so that they would be able to answer questions at the end of each passage; whereas, the control was never told to read carefully. The slowed reading times for the integrative condition was determined to be caused by taxed attentional resources. Even though the control was found to have faster reading times overall, their reading times still slowed during the inference inducing sentences, which would indicate that they were still making the inference online.

Cain and Oakhill (1999) examined the relation between comprehension and inferencing ability in 6-8 year old children while reading short narrative stories. The first two groups were comprised of children ages 7-8 that were split by high and low comprehension scores, but they were matched for reading accuracy, sight vocabulary, and chronological age. The third group was comprised of 6-7 year old children considered normally developing readers with equivalent comprehension skills to the low skilled comprehenders. They were labeled the comprehension-age matched (CAM) group and provided a strong test to the directionality hypothesis being tested. Basically, the study investigated whether good comprehension contributes to a better ability to draw inferences or does the ability to draw inferences aid comprehension. The study found that the poor comprehenders were able to answer the same amount of literal questions about the stories read, but they could not answer as many text-connecting inferential questions as their higher level peers. More importantly, the poor comprehenders were not able to answer as many text-connecting inferential questions as the CAM group. When a child gave an incorrect answer, their mistake was pointed out to them and they were given the opportunity to search through the text for the correct answer. The poor comprehenders were able to improve their performance on both literal and text-connecting inferential questions when given this assistance. This demonstrates that the poor comprehenders are capable of making text-connecting inferences given time to go back and search, but apparently they do not do so spontaneously.

A more recent study by Cain, Oakhill, Barnes, & Bryant (2001) examined the possible sources of the poor comprehenders' difficulty with making causal antecedent inferences. To explore this, 7 and 8 year old students in the study were matched according to their knowledge base by being taught a series of facts about an imaginary planet. These facts provided the background knowledge needed to make the causal antecedent inferences built into the text. The results showed that a lack of background knowledge was not the only contributing factor to the low text-connecting inferencing rate of poor comprehenders. The study proposed that a breakdown in one of three stages in the inferencing process may be the primary source of inferencing failure. The three stages include premise recall, integration, and inference formation. Premise recall is an information retrieval process where the reader pulls information from previous text or from their knowledge base that is needed to make the inference. Integration is where the information needed from previous text is brought together with knowledge base information, which leads to the third and final stage, formation of an inference. Once all of the needed information is brought together, an inference is generated by the reader to establish coherence of the text. If the reader has difficulty at any point in this sequence, it is unlikely that an appropriate inference would be made. The study revealed that a common source of inferencing difficulty for poor comprehenders was failure to retrieve the relevant information from prior text. The more skilled comprehenders were able to retrieve the relevant textual premise, but their difficulties in making inferences were traced to integrating the textual premise with the knowledge base information. It is important to note that the less skilled readers would most likely have integration difficulties themselves; however, they don't reach that stage in the inferencing process due to the initial failure to retrieve the relevant premise information.

This failure to retrieve relevant information previously presented in the text may be a cognitive limitation problem for the less skilled comprehenders. The process of retrieving relevant information may become easier as the task of decoding becomes more automatized. When considering automaticity in reading it is important to consider four properties: speed, effortlessness, autonomy, lack of conscious awareness (Logan, 1997). It is possible that more skilled readers are able to make more causal antecedent inferences because they allocate fewer of their cognitive resources to reading each individual word, freeing them to take on other tasks. Poor readers allocate all of their cognitive resources on decoding the individual words and may have few left to retrieve and integrate prior information with the proper global knowledge needed to make the necessary inference.

Although studies by Casteel and Simpson (1991) and Casteel (1993) found that second and third graders were able to generate causal inferences, they were all generated during readings of narrative text. Reading of expository text imposes many different constraints on both knowledge base and reading goal. Reading of expository text can require the use of more cognitive resources to comprehend, which may make it more difficult to generate inferences online. Also, much of the information presented in expository text is novel to the reader, which makes it more unlikely that the necessary knowledge base would be in place from which to draw an inference. Although students may try to integrate previous text with prior knowledge to maintain local coherence, as Cain, Oakhill, Barnes, and Bryant (2001) suggest, it is impossible to integrate prior text to a knowledge base that does not exist. The study presented here was designed to minimize the difficulties with prior knowledge through the use of knowledge base questions and eliminate passages for which the child does not have the necessary knowledge to make the inference.

In this study, participants were presented six-sentence expository passages. The children read pairs of sentences aloud, after which an unrelated word, an associated word, or an inference word were presented and the child read the word as quickly and accurately as possible. If the child read the inference word quicker than the unrelated word and the child passed the knowledge base question for that item, it was assumed that the child had made the inference during the reading of the passage. If the child read the associated word faster than the unrelated word, it was assumed that the child read the passage with some degree of comprehension.

In this study, I am hypothesizing that the third grade students will generate more inferences than the second grade students when reading expository text. All of the previous studies that used narrative text suggest this to be true (Casteel & Simpson, 1991; Casteel, 1993). Also, I expect high skilled readers to generate more inferences than the low skilled readers due to the fact that they have more cognitive resources to allocate to inference generation that are not being consumed by the decoding process.

CHAPTER 2: METHODS

Design

A 2 grade (second and third) X 3 target type (unassociated, associated, and inference words) mixed factorial design was used for this study. Grade was a between-subjects factor and target type was a within-subjects factor. Word naming time in milliseconds (ms) and word naming accuracy were the dependent measures.

Participants

Participants were 64 second graders and 78 third graders. Participants were from schools in an urban setting in Northeast Georgia and a suburban setting in Northern New Jersey. The 65 participants from the Georgia site were approximately 70% African American, 5% white, 20% Hispanic, and 5% Multiracial. The 77 participants from the New Jersey site were approximately 10% African American, 45% White, 30% Hispanic, and 15% Asian. All children attended Title 1 schools. None of the participating children who have received English as a Second Language services were included in the study. Schools in the Georgia site had approximately 89% of their students eligible to receive free and reduced lunch and the New Jersey schools had approximately 35% of students eligible for free and reduced lunch. Students were part of a larger study of the development of reading fluency and all received parental permission to participate.

Stimuli

Passages used for the experiment were taken from authentic expository science texts written for children reading at the second grade level. Finding “true” expository texts for second grade readers proved to be a difficult task. Selected passages were determined to be from expository text if the text they were taken from was informational in nature rather than following traditional story structure. The texts from which passages were drawn were written to provide the reader with new knowledge about the outside world rather than being written solely for entertainment value. Passages were selected as potential passages for the study because the experimenters believed that they contained segments requiring causal inferences according to the definition provided by Millis and Graesser (1994). The passage segments taken from these texts were modified to fit the six-sentence structure needed to fit our design. All trimming of the passages was done carefully so as not to alter the meaning or authenticity of the text unduly. Of the 12 passages selected, 1 was eliminated due to a typographical error, leaving 11 passages to be used in the study. (As one can see, later I needed to eliminate a second passage because target word difficulty, so what is reported below reflects the ten final passages included in the analyses.)

In order to ensure that inference words chosen for the study were not mere associates, associations were collected for all content words in the passage. A list of all words carrying meaning was derived from all potential passages to be used in the study. Twenty-five undergraduate college students enrolled in an introductory Psychology course were used to norm the passages. They were asked to read the list of 241 randomly ordered words and put down the first thing they associated with each word. The words most often listed were used as the associates for each corresponding passage. In addition these lists were used to ensure that the inference words being used were not associates in any way to words contained in the same passage. None of the potential inference words in the final passages selected were listed by any of the participants.

These same adults then read each of the passages and were asked a “why” or “because” question that would elicit the inference response that was required for understanding the passage but was not explicitly stated within the passage (hence the bridging inference). All selected inference words were listed as answers to these questions by at least 50% of the students. Of the alternative answers given, 75% of them were close associates to the desired inference word.

Finally, a list of non-associates was derived using the same list of associate words. Words that were not listed to any of the content words were considered unassociated. All three types of words were cross-examined using the second grade corpus of The Educator's Word Frequency Guide (Zeno, Iven, Millard, & Duvvuri, 1995) to ensure that there was no significant difference in the frequencies of the three different conditions at a second grade level criterion (associated, $M= 193.8$, $SD= 61.6$; inference, $M=195.6$, $SD= 70.4$; unassociated, $M=196.1$, $SD= 64.1$). Target words were also controlled for word length in number of letters (associated, $M= 4.3$, $SD= 1.4$; inference, $M= 5.0$, $SD= 1.5$; unassociated, $M= 4.6$, $SD= .7$). Three t-tests confirmed that there were no significant differences is word length between the three respective target types (inference/ associate: $t(8)= 13.286$. $p= .76$, inference/ unassociated: $t(8) = 24.00$, $p= .76$, associate/ unassociated: $t(8) = 29.667$, $p= .76$).

Procedure

E-Prime software (Psychology Software Tools, 1999) was used to present stimuli and record reaction times. Stimuli were presented on a Dell 8100 laptop and responses were recorded using a Audio-Technica ATR 20 microphone and PST Response Box. Participants were tested in a quiet area in the child's school.

Each passage contained six related sentences. Each six-sentence passage was divided into three subsections that contained two sentences each. Each of the subsections was paired with either an inference word, an associate, or a non-associate. Each passage contained one of each of these target word types. The order in which the stimulus words appeared was controlled by alternating the order they were presented in each passage. For example, the inference word appeared in each position (1st, 2nd, and 3rd) an equal number of times to control for priming effects. If necessary, the passage was re-written slightly so that the passages were coherent but still permitted counterbalancing the order of the target word conditions.

A typical trial began with a ready signal followed by a focus point in the middle of the screen for 500ms. Then the first two sentences of the passage would appear. Once the student had attempted the first two sentences the experimenter would rate their reading on the response box by one of three categories. It was rated as a "one" if the sentences were read correctly, it was rated a "two" if a few words were misread but the meaning of the sentences was not altered, and it was rated a "three" if the sentences were read incorrectly and the meaning was changed by the reader's mistakes. Once the experimenter rated the student's reading, the sentences would disappear from the screen and another focal point would appear in the center of the screen for another 500ms. This focal point would be followed by a target word that would appear in the middle of the screen until the child made an attempt to read it. The microphone would trigger the word to disappear and the experimenter would rate the child's reading of the target word. The experimenter would rate it a "one" for a correct response, a "two" for an incorrect response, and a "three" if environmental noise or some other experiment-irrelevant issue caused the microphone to trigger before the child was able to read the word. This process was repeated for two practice trials and three to twelve experimental trials for each participant. The trials were terminated if after three sequential trials the child was unable to read the passages fluently enough or the target words accurately.

After the child completed the experimental task he or she was asked a series of questions aimed at determining whether the child could draw the required inference when explicitly asked to do so. Any item for which the child was unable to answer the explicit inference question was eliminated from further analyses. An example of a passage and its corresponding explicit inference question can be found in Table 2.1. Participants were also given the Wechsler Individual Achievement Test – Reading Comprehension Subtest at a separate session. The scores obtained from the Wechsler Individual Achievement Test- Reading Comprehension Subtest were used to determine reading skill. The ranges of standard scores within each skill level are presented in Table 2.2.

Table 2.1

Example Passage and its Corresponding Inference Eliciting Question

Passage	Target
Birds have beaks.	
A bird's beak helps it get food and eat it.	Fly (associate)
Parrots have short, sharp beaks.	
They use their beaks to open hard nuts.	Crack (inference)
Robins have small, pointed beaks.	
They use their beaks to catch worms.	Stone (unassociated)

Inference Eliciting Question
How would a bird with a short, sharp beak open a hard nut?
a) They would throw it on the ground.
b) *They would use their beak to crack it.*
c) They would have a friend open it.

Table 2.2

Standard Scores from the WIAT Used to Determine Reading Skill Level

Skill Level	Standard Score Range
1 (High)	104-151
2 (Middle)	95-103
3 (Low)	67-94

CHAPTER 3: RESULTS

A multi-step procedure for eliminating non-valid naming times was used. All target naming times for trials where participants were deemed disfluent in the reading of the passage were eliminated. Further, all incorrect naming times were eliminated and an error was scored. Next, all naming times where the microphone was influenced by outside noise or some other problem unrelated to the experiment were eliminated for each student (13.3 % of trials). All naming times for passages where the participant was not able to make the explicit inference when asked were eliminated (7.0% of trials). Next, the mean reaction times and standard deviations for target trials were calculated from the remaining target naming times. Any times that were outside of 2 standard deviations from their individual mean were eliminated (4.02% of trials). Finally, one of the stimulus passages was eliminated due to a typographical error, it was missing the inference target; therefore, all items from that passage were eliminated. Further, another passage proved to be too difficult for the students as noted by the high error rates (>25% trials) for the inference target, so it was also eliminated. The relatively low number of passages caused relatively low power in the items analysis; therefore, the subject's analysis will be the main analysis for discussion in this study. The results of the items analysis can be found in Appendix B. For determining reading skill, children's scores on the Wechsler Individual Achievement Test –Reading Comprehension subtest were used. Children were divided into 3 groups on the basis of this subtest. The means as a function of age, skill, and target type can be found in Table 3.1.

Table 3.1

Subject Means for Unassociated, Associated, and Inference Target Words Split by Grade and Skill Level

Grade		Skill	Unassociated	Associated	Inference
2	Naming Time (ms)	Hi (1)	901	842	870
		Md (2)	888	769	940
		Lo (3)	1214	1067	1182
	Naming Accuracy	Hi (1)	.937	.984	.916
		Md (2)	.971	.967	.874
		Lo (3)	.885	.927	.678
3	Naming Time (ms)	Hi (1)	650	627	651
		Md (2)	749	685	769
		Lo (3)	796	749	850
	Naming Accuracy	Hi (1)	.988	.992	.978
		Md (2)	.988	.982	.950
		Lo (3)	.938	.948	.806

<u>Naming Time Analysis</u>

I conducted a 2 (Grade) X 3 (Target Type) ANOVA using reaction time as the dependent variable. The third grade readers were faster at recognizing words than the second grade readers, $F(1, 140) = 17.82$, $p<.001$. The main effect of target type indicated that associate words benefited from context more than unassociated and inference words did, $F(2, 280) = 10.32$, $p<.001$. This effect was similar across grade, Target Type X Grade $F(2, 280) = 1.07$, $p=.343$. The lack of a Target Type X Grade interaction clearly indicates that there was no inferential priming regardless of grade level.

Follow-up partial ANOVAS were conducted using two separate 2 (Grade)X 2 (Target Type) ANOVAS using reaction time as the dependent variable comparing both associated and inference reaction times with the unassociated reaction times. The analyses show that the associated words benefited from contextual priming by comparing reaction times of the associated and unassociated words, $F (1, 143) = 22.08$, $p < .001$. This effect was not similar across Grade, Target Type X Grade $F (1, 143) = 4.58$, $p = .034$. This weak interaction indicates that the second grade readers benefited from associate priming slightly more than their third grade peers. A contrast of the inference reaction times with the unassociated reaction times showed no indication of contextual priming effects, $F (1, 140) = .23$, $p = .633$. This effect remained constant over grade, Target Type X Grade, $F (1, 140) = .364$, $p = .547$. This strongly indicates that inferential priming was not evident regardless of grade level.

To assess whether the highest skill readers might show more inferencing than their low skill counterparts, I added in Skill as a factor in the analysis and conducted a 2 (Grade) X 3 (Skill) X 3 (Target Type) ANOVA using reaction time as the dependent variable. High skill readers were faster at recognizing words than low and medium skill readers, $F (2, 136) = 6.33$, $p = .002$. Third grade students read the words faster than second grade students, $F (1, 136) = 20.34$, $p < .001$. The main effect of target type indicated that associated words benefited from context more than unassociated and inference words did, $F (2, 272) = 10.24$, $p < .001$. This effect was similar across Grade, Target Type X Grade $F (2, 272) = 1.26$, $p = .326$; and Skill, Target Type X Skill $F (4, 272) = 1.15$, $p = .334$; Target Type X Grade X Skill $F (4, 272) = .41$, $p = .80$. The Target Type X Grade, Target Type X Skill, and Target Type X Grade X Skill interactions are important because they tell us if the older and/or more skilled readers were making more inferences than their younger/ less skilled peers. In this case the results clearly show that inferences were not being made regardless of grade or skill level.

Accuracy Analysis

To assess whether effects of inferencing might manifest itself in greater naming accuracy if not speed, I conducted a 2 (Grade) X 3 (Target Type) ANOVA using accuracy as the dependent variable. Third grade readers were more accurate at identifying target words than second grade readers, $F(1, 143) = 7.27$, $p=.008$. The main effect of Target Type indicated that associated words were read more accurately than unassociated and inference words, $F(2, 286) = 40.51$, $p<.001$. However, this effect was not similar across grade, Target Type X Grade $F(2, 286) = 4.74$, $p=.009$. The Target Type X Grade interaction indicates that the effect of Target Type diminished as the students got older.

Follow-up partial ANOVAS were conducted using two separate 2 (Grade) X 2 (Target Type) contrasts with accuracy as the dependent measure comparing both associated and inference with unassociated accuracies. The analysis of associated and unassociated accuracies shows a weak main effect of grade, $F(1, 143) = 4.26$, $p=.041$, which indicates that the second graders were less accurate than third graders. However, there was no main effect of Target Type, $F(1, 143) = 3.56$, $p=.061$, or Target Type X Grade interaction, $F(1, 143) = 2.43$, $p=.121$. However, given that the naming accuracies in these conditions are at ceiling, it is not surprising that no beneficial effects of associate priming were found. The contrast comparing inference words with unassociated words showed a main effect of Target Type; however, the relationship is in the opposite direction indicating no beneficial inferential priming effects. The second grade students were less accurate than the third grade students, $F(1, 143) = 7.72$, $p=.006$. There was also a main effect of Target Type, $F(1, 143) = 47.75$, $p<.001$, and this effect remained consistent across grade, Target Type X Grade $F(1, 143) = 3.71$, $p=.056$ but was in the direction of second grade students finding the inference words more difficult than third grade students. These results would indicate that the inference words were more difficult for all students regardless of grade level.

As with naming times, I tested whether high skill readers might show beneficial priming in terms of naming accuracy. I added in Skill by conducting a 2 (Grade) X 3 (Skill) X 3 (Target Type) ANOVA using accuracy as the dependent variable. High skill readers were more accurate than low and medium skill readers, $F(2, 139) = 16.69$, $p<.001$. Third grade students were also more accurate in naming targets than second grade students, $F(1, 139) = 9.08$, $p=.003$. The main effect of Target Type seemed to indicate that inference words were read less accurately than associated and unassociated words, $F(2, 278) = 45.49$, $p<.001$. However, this effect was larger for second graders than third graders, Target Type X Grade $F(2, 278) = 5.60$, $p=.004$; and particularly large for low skilled readers, Target Type X Skill $F(2, 278) =11.67$, $p<.001$. However, the three-way interaction between Target Type X Grade X Skill proved non-significant, $F(2, 278) = .51$, $p=.730$. The significant Target Type X Grade and Target Type X Skill interactions reflect that the differences in accuracy among target type conditions diminished somewhat as the readers got older and more skilled.

CHAPTER 4: DISCUSSION

The purpose of the present study was to investigate whether students were making inferences while reading expository texts at various age and skill levels. The results of the present study did not find evidence that any of the students were making inferences during the reading of expository texts regardless of grade or skill level. In fact, the results indicate that the students were having a more difficult time with the inference target words than with the associated and unassociated target words. Noordman, Vonk, and Kempff (1992) note that longer reading times may be due to extra time needed to retrieve past information and make the connection with the text when formulating a causal inference. If anything, this suggests that students were making inferences after they read the target words and tried to integrate these into the representation of the passage. Another possible reason for students taking longer to name the inference target words might have been that students mentally generated synonyms for the inference words rather than the actual inference words themselves because of the time lag between the offset of the passage and the onset of the stimulus word. That is, students might have been making the inference online; however, the inference active at the time of reading could have been a close synonym of the desired target word. The extra naming time could reflect the time it took the child to make the connection between the active synonym and the desired target word. The difficulty with inference words was not consistent over grade level and seemed to dissipate as the students got older and more skilled as readers; however, as expected, students from all age and skill groups read the associated words quicker and more accurately than the unassociated words.

Previous research of young children's ability to make causal inferences while reading have mainly focused on narrative text rather than expository text. In these studies (Casteel, 1993; Casteel & Simpson, 1991), students as young as second grade were found to draw inferences while reading narrative texts. There are many differences between these studies with my own that may explain why these students were drawing inferences in one situation but not in the other. In both Casteel studies the students were selected as "average readers" from an "affluent school district". This study dealt with all readers from three economically disadvantaged schools. For this reason, the second and third grade students in both Casteel studies were most likely at an entirely different reading level than the students who participated in this study. Also, the fact that students at this age have more experience with narrative text rather than expository text may influence their ability to draw inferences while reading each type of passage.

Past research that has examined people's ability to draw inferences while reading expository text has focused almost exclusively on adults. This design was modeled after the methods used by Millis and Graesser (1994); however, it needed to be altered slightly in order to be implemented with children rather than the adults used in the Millis and Graesser (1994) study. In order to ensure that these young students would have the stamina to remain focused and give their best effort throughout, the test was given using 12 passages rather than the 44 passages used in the Millis and Graesser (1994) study. Selecting developmentally appropriate passages was also made difficult due to the wide range of reading skill levels present in second and third grade readers. Students range from young adult levels down to emerging readers. Finding passages that all students would be able to attempt but also would stay below ceiling for most third grade readers was a difficult task.

The study by Millis and Graesser (1994) along with ones by Singer, Harkness, & Stewart (1997); Narvaez, van den Broek, & Ruiz (1999); and Noordman, Vonk, & Kempff (1992) all indicate that adults are able to draw inferences while reading expository text. This indicates that the ability to draw such inferences develops by adulthood, but little has been done to pinpoint exactly when this skill develops during reading development. Although the current study does not pinpoint exactly when students begin to draw causal inferences while reading expository text, it does begin to narrow the age range necessary to be included in future research on this topic.

Studies by Cain & Oakhill (1999) and Cain, Oakhill, Barnes, & Bryant (2001) have attempted to pinpoint why some children are able to draw inferences and others are not. These studies found that comprehension skill and background knowledge played a key role in students' ability to draw inferences while reading. This study was designed to eliminate the issue of background knowledge through the use of explicit inference questions related to each individual test passage following the experimental task. If the information necessary to draw the explicit inference was unavailable to the student, I eliminated all accuracies and naming times for that particular passage. The issue of comprehension skill was taken into account through the use of the Reading Comprehension Subtest of the Wechsler Individual Achievement Test and allowed me to break the students into separate reading skill groups. However, even the highest skilled readers from this group were most likely at a disadvantage to the skill groups used in both Oakhill (1999, 2001) studies.

Future research in this area should be directed at adults' ability to draw inferences while reading expository texts. Past research by Millis and Graesser (1994); Singer, Harkness, & Stewart (1997); Narvaez, van den Broek, & Ruiz (1999); and Noordman, Vonk, & Kempff (1992) have all shown that adults are able to draw inferences when reading expository text; however, each of these studies has used pools of students from undergraduate courses at the University level. This population of subjects does not represent an average adult population. It is possible that the average or below average adult reader may not draw these inferences while reading, in which case it would be useless to continue testing children from this population to determine when this inferencing ability develops. If it is found that these adults do make inferences when reading, the next logical step would be to determine when this skill develops.

This study did not find evidence that second and third grade readers from low SES backgrounds make inferences when reading expository texts, regardless of skill level. Although the results failed to show that the third grade students were making inferences, there were some slight indications that they were not far off. The next step would be to begin testing students from third, fourth, and fifth grades using refined methods to examine if elementary aged kids develop the ability to draw causal inferences while reading expository texts at all. There is a switch from "learning to read" to "reading to learn" when students jump from third grade to fourth grade. It could be possible that the extra emphasis on reading to learn from a variety of expository texts may give these students the experiences with expository text necessary to foster the development of their ability to draw these causal inferences.

The developmental level of the readers tested in this group caused many of the challenges faced in the design of this study. Testing adults and older readers would not present many of these challenges and would allow for better stimuli to be selected. Causal inferences require a certain level of syntactic complexity that is difficult to find in texts designed for emerging readers. Students in the intermediate grades are at a skill level that would make the selection of valid stimuli a much easier task.

Although the present study did not discover the exact age or reading level at which inferences are generated when children are reading expository texts, it did provide valuable information to be used in future research. The data from this study failed to indicate that second and third grade students regardless of age or reading level were generating causal inferences. Future research needs to examine whether adults make such these inferences. Based on the results of this study we can be reasonably sure that this does not occur in a manner that benefits the processing of inference words until after the third grade.

REFERENCES

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing: An Interdisciplinary Journal*, 11, 489-503.

Cain, K., Oakhill, . V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory and Cognition*, 29, 850-859.

Casteel, M. A. (1993). Effects of inference necessity and reading goal on children's inferential generation. *Developmental Psychology*, 29, 346-357.

Casteel, M. A., & Simpson, G. B. (1991). Textual coherence and the development of inferential generation skills. *Journal of Research in Reading*, 14, 116-129.

Graesser, A. C., Swamer, S. S., Baggett, W. B., & Sell, M. A. (1996). New models of deep comprehension. In B. K. Britton & A. C. Graesser (Eds.), *Models of Understanding Text* (pp.1-31). Mahwah, NJ: Erlbaum.

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly: Overcoming Learning Difficulties,* 13, 123-146.

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

Millis, K. K., & Graesser, A. C. (1994). The time-course of constructing knowledge-based inferences for scientific texts. *Journal of Memory and Language*, 33, 583-599.

Narvaez, D., van den Broek, P., & Ruiz, A. B. (1999). The influence of reading purpose on inference generation and comprehension in reading. *Journal of Educational Psychology*, 91, 488-496.

Noordman, L. G., Vonk, W., & Kempff, H. J. (1992). Causal inferences during the reading of expository texts. *Journal of Memory and Language*, 31, 573-590.

Schneider, W. (1999). E-Prime [Computer software and manual]. Pittsburgh, PA: Psychology Software Tools Inc.

Singer, M., Harkness, D., & Stewart, S. T. (1997). Constructing inferences in expository text comprehension. *Discourse Processes*, 24, 199-228.

Trabasso, T., van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, 24, 612-630.

Zeno, S. M., Iven, S. H., Millard, R. T., & Duvvuri, R. (1995). The educator’s word frequency guide. New York, NY: Touchstone Applied Science.

APPENDIX A: TEST ITEMS AND TARGETS

Every dolphin is different.	
Some are shy.	Ocean
Some show off.	
Some like to tease.	Grass
One dolphin liked to yank the tail feathers of a pelican.	
The pelican didn't think it was funny.	Hurt
Dolphins can be smart.	
A dolphin named Kathy liked to play with a small rubber ring.	Line
One day a man came to watch her.	
Kathy flipped her ring to him.	Night
Two years later, the same man came back.	
The minute Kathy saw him she tossed her ring right to him.	Remember
A baby whale cannot swim very fast.	
A big, hungry shark is watching.	Real
It is waiting for a chance to attack.	
When his mother sees the shark, she rushes straight at it.	Protect
Mother whales can swim fast.	
The baby whale is safe again.	Father
Sea otters live and sleep in the sea.	
They just close their eyes and go to sleep on their backs.	Float
The water where the sea otters live is cold.	
But the sea otters don't mind.	Ice
They love to play in the waves.	
They race and chase each other.	Sick

The noble lion lets his wife do most of the work.	
He does help her in one way.	Husband
He roars.	
The other animals run away.	River
They run right toward the lioness.	
Then he joins her for dinner.	Kill

Grandfather gave me a glass jar to take on our firefly hunts.	
We punched holes in the lid.	Air
Fireflies look like dancing stars.	
They are really beetles, grandfather says.	Bug
Their front wings are hard and their back wings are soft.	
They live in the ground most of the time.	Number

Even small bugs can be strong enough to fight their enemies.	
Some of them can run or jump quickly, while others can fly away.	Escape
A bug may be able to sting enemies, or have strong jaws to bite them.	
Even so, lots of bugs are killed.	Bee
But there are always more.	
Bugs are born by the thousands.	Road

Winter is coming.	
The days grow short.	Tall
The nights grow long.	
Leaves have fallen from the trees.	Walk
There are no berries on the bushes.	
The grass is brown.	Die

Some spiders spin tangled webs.	
When an insect is trapped, the spiders run to get it.	Food
Other spiders weave sheet webs.	
The spider hangs upside down under the web.	Black
When an insect hits the sheet web, the spider quickly pulls it through.	
Spiders may look scary but most of them don't hurt people.	Fast

Birds have beaks.	
A bird's beak helps it get food and eat it.	Fly
Parrots have short, sharp beaks.	
They use their beaks to open hard nuts.	Crack
Robins have small, pointed beaks.	
They use their beaks to catch worms.	Stone

APPENDIX B: ITEMS ANALYSIS

Naming Time Analysis

Using item as the random variable, I conducted a 2 (Grade) X 3 (Target Type) X 3 (Skill) ANOVA with reaction time as the dependent variable. The third grade readers were faster at recognizing words than the second grade readers, $F(1, 9) = 130.78$, $p<.001$. The main effect of target type was lost due to the small number of items in the analysis, $F(2, 18) = .671$, $p = .524$. The lack of an effect was similar across grade and skill, Target Type X Grade $F(2, 18) = .128$, $p = .88$, Target Type X Skill $F(4, 36) = .226$, $p = .922$, and Target Type X Grade X Skill $F(4, 36) = .209$, $p = .932$. Again, the lack of any interaction may be attributed to the low power of the analysis caused by the small number of items analyzed.

Accuracy Analysis

To assess whether effects of inferencing might manifest itself in greater naming accuracy if not speed, I conducted a 2 (Grade) X 3 (Target Type) X 3 (Skill) ANOVA using accuracy as the dependent variable and item as the random variable. The third graders were more accurate at naming words than the second grade readers, $F(1, 9) = 13.84$, $p = .005$. The main effect of Target Type indicated that associated words were read more accurately than the unassociated and inference words, $F(2, 28) = 4.08$, $p = .034$. This effect was similar across skill, Target Type X Skill $F(4, 36) = 3.91$, $p = .01$. The Target Type X Skill interaction indicates that the effect diminished as the students got more skilled as readers. However, this effect was not similar across grade, Target

Type X Grade F (2, 18) = 3.22, p = .064, Target Type X Grade X Skill F (4, 36) = 1.29, p = .290. The lack of a Target Type X Grade and Target Type X Grade X Skill interaction indicates that the effect remained constant as the students got older.