

THE ROLE OF QUESTIONING IN CREATING SITUATION MODELS WHILE READING
IN A SECOND LANGUAGE: DOES EXPLAINING EVENTS IN A TEXT MATTER?

by

MIKI LOSCHKY

B.A., Stephens College, 1981
M.A., University of Hawaii, 1989

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Curriculum and Instruction
College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2014

Abstract

The primary purpose of this study was to explore ways in which teachers can increase their second language (L2) learners' reading comprehension through constructing situation models. The author incorporated theoretical frameworks, including the situation model theory (Kintsch, 1998/2007), the event index model (Zwaan & Radvansky, 1998), and the linguistic threshold hypothesis (Clark, 1980; Aldersen, 1984; Carrell, 1991). As an educational intervention, a set of adjunct questions were asked during reading to elicit readers' explanations of causality and intentionality to promote coherent comprehension. A total of 117 L2 readers' ability to make appropriate inferences based on situation models was assessed through both quantitative (experimental-control design) and qualitative (think-aloud) methods. This experimental study examined the effect of explaining to target situation models, while also looking at the relationship between the readers' inferential ability and their L2 proficiency. In addition to the main effects of the intervention and L2 proficiency, the interaction between the intervention and L2 proficiency were discussed as results of the quantitative analysis. Also discussed was the nature of the L2 readers' responses to adjunct questions, which were designed to measure the quality of their explanations based on the underlying situations in the text that they were reading.

This study extended existing research on situation model-level comprehension to L2 literacy, which has not previously been well studied. This made the study theoretically interesting as well as highly applicable to L2 reading instruction. The main findings of this study were: (1) there was a strong effect of question types (inferential vs. non-inferential questions) with greater accuracy overall for non-inferential questions, (2) there was a suggestive trend of question type interacting with L2 proficiency, such that higher proficiency participants showed little difference between question types, (3) there was no statistically significant main effect of the adjunct question manipulation on accuracy, and (4) there was a trend suggesting an interaction between experimental condition and L2 proficiency, with higher proficiency participants showing a somewhat larger effect of the adjunct question manipulation. Additionally, both quantitative and qualitative data trended in the direction consistent with the linguistic threshold hypothesis.

THE ROLE OF QUESTIONING IN CREATING SITUATION MODELS WHILE READING
IN A SECOND LANGUAGE: DOES EXPLAINING EVENTS IN A TEXT MATTER?

by

MIKI LOSCHKY

B.A., Stephens College, 1981

M.A., University of Hawaii, 1989

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Curriculum and Instruction
College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2014

Approved by:

Major Professor
Dr. Socorro Herrera

Copyright

MIKI LOSCHKY

2014

Abstract

The primary purpose of this study was to explore ways in which teachers can increase their second language (L2) learners' reading comprehension through constructing situation models. The author incorporated theoretical frameworks, including the situation model theory (Kintsch, 1998/2007), the event index model (Zwaan & Radvansky, 1998), and the linguistic threshold hypothesis (Clark, 1980; Aldersen, 1984; Carrell, 1991). As an educational intervention, a set of adjunct questions were asked during reading to elicit readers' explanations of causality and intentionality to promote coherent comprehension. A total of 117 L2 readers' ability to make appropriate inferences based on situation models was assessed through both quantitative (experimental-control design) and qualitative (think-aloud) methods. This experimental study examined the effect of explaining to target situation models, while also looking at the relationship between the readers' inferential ability and their L2 proficiency. In addition to the main effects of the intervention and L2 proficiency, the interaction between the intervention and L2 proficiency were discussed as results of the quantitative analysis. Also discussed was the nature of the L2 readers' responses to adjunct questions, which were designed to measure the quality of their explanations based on the underlying situations in the text that they were reading.

This study extended existing research on situation model-level comprehension to L2 literacy, which has not previously been well studied. This made the study theoretically interesting as well as highly applicable to L2 reading instruction. The main findings of this study were: (1) there was a strong effect of question types (inferential vs. non-inferential questions) with greater accuracy overall for non-inferential questions, (2) there was a suggestive trend of question type interacting with L2 proficiency, such that higher proficiency participants showed little difference between question types, (3) there was no statistically significant main effect of the adjunct question manipulation on accuracy, and (4) there was a trend suggesting an interaction between experimental condition and L2 proficiency, with higher proficiency participants showing a somewhat larger effect of the adjunct question manipulation. Additionally, both quantitative and qualitative data trended in the direction consistent with the linguistic threshold hypothesis.

Table of Contents

List of Figures	viii
List of Tables	ix
Acknowledgements.....	x
Chapter 1 - Introduction.....	1
Overview of the Issues.....	1
Theoretical Framework.....	4
Statement of the Problem.....	7
Purpose of the Study/Research Questions	8
Significance of the Study	8
Limitations of the Study	9
Definition of Terms	10
Chapter 2 - Review of the Literature	13
Cognitive Processing and General Domain Knowledge.....	14
Different Levels of Text Representations.....	17
Top-down and Bottom-up Processing	20
Inference Generation.....	24
The Role of Working Memory in Drawing Inferences.....	25
Coherence vs. Cohesion.....	28
Implications of Inference Drawing Abilities for ELL Students	29
Interaction with Text.....	31
Questioning and Explaining as a Strategy to Create Situation Models during Reading	32
Knowledge Base vs. Language Proficiency	35
Summary.....	36
Chapter 3 - Methodology	38
Research Questions.....	38
Research Design	39
Population and Sample	40
Data Collection	41

Materials	41
Procedure	41
Data Analysis	42
Quantitative measures – Reading comprehension task (paper and pencil).....	43
Qualitative measures - Think-aloud/verbal protocols.....	43
Reliability and validity.....	45
Chapter 4 - Data Results and Analysis	49
Description of the Sample Population	51
Description of the Measures	53
Analysis of Quantitative Data.....	59
Qualitative Analysis of Readers’ Thoughts (Think Aloud).....	65
Overarching ideas	78
Pragmatic knowledge.....	80
Coherence	82
Chapter 5 - Discussion, Implications, and Conclusions	85
Summary of the Findings.....	86
Theoretical and Practical Implications of the Study.....	92
Recommendations for Future Research.....	95
Conclusions.....	97
References.....	100
Appendix A - Inference Passages for the Control Condition	109
Appendix B - Inference Passages for the Experimental Condition	114

List of Figures

Figure 4.1 Comprehension Accuracy across Question Type and L2 Proficiency	62
Figure 4.2 Comprehension Accuracy Summary across Conditions and L2 Proficiency.....	65
Figure 4.3 Level of Understanding Causal/Intentional Dimension across L2 Proficiency	71
Figure 4.4 Correlation between EPT Scores and Content Scores.....	77

List of Tables

Table 4.1 Age of Sample Population	51
Table 4.2 Proficiency of Sample Population	52
Table 4.3 Coding for Qualitative Data.....	56
Table 4.4 Summary of the 2 x 2 x 3 Mixed ANOVA (Test if Within-Subjects Effects and Within-Subjects x Between-Subjects Effects) Coding for Qualitative Data.....	59
Table 4.5 Summary of Descriptive Statistics for Comprehension Accuracy as a Function of Intervention Condition and L2 Proficiency	63
Table 4.6 Summary of the 2 x 3 ANOVA (Tests of Between-Subjects Effects).....	63
Table 4.7 Raw Frequencies of Think Aloud Responses	68
Table 4.8 Percentage of Frequencies of Think Aloud Responses	69
Table 4.9 Top 4 Think Aloud Responses across L2 Proficiency.....	72
Table 4.10 Sample Items for Content Analysis Scoring of Think Aloud Responses	75

Acknowledgements

I would like to thank my internal committee (Dr. Socorro Herrera, Dr. Richard Harris, Dr. Laurie Curtis, and Dr. Kevin Murry) and Dr. Joseph Magliano for their helpful comments and discussions of this dissertation research. Special thanks to Dr. Herrera for helping me access over 100 participants, which made my experimental research possible. I would also like to thank Andrew Dalton and Christopher Bellinger for their generosity in co-rating my qualitative data, which must have consumed a tremendous amount of their time. I appreciate the intra-departmental support from Dr. Christopher Vahl, Xia Jang, and Stuart Miller for sharing their statistical knowledge to help organize my data. Melissa Holmes and Christopher Remple get my highest respect for the thorough editing skills that they have provided in proofreading the previous version of this manuscript. Finally, my deepest appreciation goes to my partner Dr. Lester Loschky for guiding me with his in-depth knowledge of how to conduct experimental research throughout the process of this study.

Chapter 1 - Introduction

In order to truly process a text, readers need to integrate what is explicitly stated in the text with their knowledge base (i.e., create a mental representation of the situation). This type of mental model is referred to as a situation model, defined by Zwaan and Singer (2003) as “a mental representation of the state of affairs denoted by a text” (p. 85). This study examines the role of questioning in forming situation models to facilitate reading comprehension among second language (L2) readers. Discussion in this chapter is organized by the following sections: (1) overview of the issues, (2) theoretical framework, (3) statement of the problem, (4) purpose of the study, (5) significance of the study, (6) limitations of the study, and (7) definition of terms.

Overview of the Issues

In reading, students often infer to integrate new information with the old. Successful readers constantly make inferences through generating situation models regarding what the text will be about, while making use of contextual information and/or general knowledge to fill in gaps, since writers do not necessarily spell out every detail in a given text. If educators can assist their students in this area, they can foster their students’ reading comprehension with unfamiliar reading content. This study reviewed previous studies of the types of inferences successful readers or high proficient L2 readers generally draw, as opposed to struggling readers or low proficient L2 readers. It also reviewed research on how reading teachers can elicit inferences from their students as a strategy to increase coherent comprehension, especially for L2 learners. Thus, the current study focused on three key issues: (1) English language learner (ELL) students’ ability to make appropriate inferences in their L2 reading based on situation models, (2) the relationship between their L2 proficiency and their ability to make inferences, and (3) the effect

of explaining on target situation models (e.g., by answering explanatory questions about why/how a certain event occurred in the text as well as what the event caused.) The inference making patterns by students at different proficiency levels in L2 observed in this study were compared with those of previous studies. The study also examined the role of adjunct questions (e.g., why and how questions) as an educational intervention to elicit causality and intentionality in creating situation models to facilitate reading comprehension.

As various text and discourse comprehension researchers have claimed, situation models play a prominent role in language comprehension and one way to assess that is via readers' ability to draw inferences (Graesser, Singer, & Trabasso, 1994; Kintsch, 1998/2007; van den Broek, 1994; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998; Zwaan & Rapp, 2006). Comprehending texts requires more than just being able to read words phonetically, knowing the meaning of each word, and interpreting the meaning of the propositions (or the idea units containing predicates) from what is explicitly stated in the text as surface level and textbase representations (Graesser, et al., 1994). It also includes being aware of information that the writer left unsaid by drawing inferences, which involves constructing a coherent mental model to make sense of a given text at a deeper level. Various text researchers have claimed that such a coherent understanding of the text content, known as a situation model, is a prerequisite to a deeper level of understanding. Grabe (2009) defined reading comprehension as "the combination of text input, appropriate cognitive processes, and the information that we already know" (p. 74). Due to these multiple levels of representations, the text is maintained "in memory as a coherent structure, rather than as a disjointed assembly of individual pieces of information" (van den Broek, 1994, p. 539). In this study, I reviewed current theories/critical concepts of reading comprehension (including the situation model, top-down vs. bottom-up processing, coherence vs.

cohesion, and working memory) in relation to the cognitive process of drawing inferences to construct meaning. As making inferences should play a prominent role in reading comprehension for all learners (including L2 learners), helping them to be able to draw appropriate inferences through creating situation models and learning how they can apply this skill in academic settings are the key elements for their school success, I explored the ways that theories/concepts apply to L2 reading in terms of drawing inferences as a strategy to increase ELL students' comprehension. I also focused on ways in which educators can elicit students' ability to draw appropriate inferences and how students can use that ability as a strategy to improve their comprehension in English. I discussed such concepts as creating situation models through questioning and explaining, while exploring the roles of L2 language proficiency.

Readers use linguistic cues, both lexical and grammatical, to construct their mental representation of a text in conjunction with their experiential knowledge (Zwaan & Radvansky, 1998; Zwaan & Rapp, 2006). However, as ELL students have less lexical and grammatical knowledge in English compared to native English-speaking students, they may fail to comprehend the message that the writer is trying to communicate, especially when it is only implied. For example, the sentence "John dropped a vase; therefore, he had to replace it" does not explicitly state that John broke the vase. The readers need to put the two separate propositions (proposition one: dropping a vase, proposition two: replacing the vase) together to make the story coherent by generating information from their knowledge of the world (e.g., causal relationships). It is critical for all students to maintain coherence while reading a text, but it is particularly crucial for ELL students who need to process information, which is most likely to be incomplete or ambiguous, in their L2. However, L2 readers may not be able to make such inter-sentential connections due to a lack of sufficient understanding of the information

represented in those sentences and/or the functions of logical connectives such as “therefore” (Goldman & Murray, 1992).

Theoretical Framework

In order to fill in linguistic gaps, L2 readers are more likely to rely on constructing “a coherent representation of a text by utilizing information explicitly presented in the text and information generated from general background knowledge” (Horiba, 1996, p. 437). This is where situation models (van Dijk & Kintsch, 1983) come into play. As defined earlier, a situation model is a state of affairs or a particular situation being described in a text (Zwaan & Singer, 2003). According to van Dijk & Kintsch (1983), such mental representations are not necessarily stored in memory as textual forms, but they are integrated with the individual’s knowledge structures during the course of reading. The following example illustrates how such knowledge structures are created in the reader’s mind:

1. Three turtles rested on a floating log, and a fish swam beneath them.
2. Three turtles rested beside a floating log, and a fish swam beneath them.

(Bransford, Barclay, & Franks, 1972 in Carroll, 2008, p. 172).

The key idea of “A fish swam beneath a floating log” is more comprehensible to readers who have read “Three turtles rested on a floating log, and a fish swam beneath them” rather than those who read in “Three turtles rested beside a floating log, and a fish swam beneath them,” because the comprehender creates a mental representation of the situation that reflects spatial relationships among things described in the sentence (e.g., the turtles, log, fish). To create a coherent situation model of the text, L2 readers must apply their knowledge of the world to make sense of ambiguous propositions.

A cognitive psychologist, Walter Kintsch (1998/2007) theorizes that inferences are a means to fill knowledge gaps, as high-knowledge readers are more likely to generate appropriate inferences when reading low-coherence text, compared to low-knowledge readers. In other words, skilled readers must have developed their ability to represent the meaning of the text coherently (i.e., in a situation model). Zwaan and Radvansky (1998) provide a thorough explanation of the role of situation models in integrating mental representations of what is being read. In their analysis situation models and inferences are separate but interrelated theoretical constructs. According to situation model theory, the reader would create a situation for whatever the text is about, and “incoming information would be linked to this token based on grammatical and world knowledge” (p. 164). Observe the example sentences that Zwaan and Radvansky provide:

Lamar Alexander was behind in the polls. However, the former Tennessee governor remained optimistic. He considered it likely that a moderate candidate with new ideas would win the Republican nomination (1998, p. 163).

The reader’s grammatical knowledge allows him or her to understand that “he” in the third sentence refers to Lamar Alexander in the first sentence. However, to infer that Lamar Alexander (sentence one), the former governor of Tennessee (sentence two), and a moderate candidate with new ideas (sentence three) refer to the same person, it is crucial for the reader to know who Lamar Alexander is. In other words, to make the appropriate inference the reader needs to know that Lamar Alexander is the former moderate Republican governor of Tennessee who ran for President, as none of the sentences explicitly says so. This particular knowledge about politics permits the reader to make sense of the ambiguous passage above.

Situation models can provide information links to various narrative dimensions, including time, space, causality, intentionality, and protagonists (Magliano, Zwaan, & Graesser, 1999; Zwaan & Brown, 1996; Zwaan et al., 1995; Zwaan & Radvansky, 1998), which are also likely to change continually during the course of reading. In other words, as readers are actively engaged in text comprehension, they construct coherent situation models as memory representations by indexing when, where, and why an event occurred as well as what the goal was and whose action took place. Zwaan and colleagues proposed and referred to this theory of multi-dimensionality within the situation model as the event-indexing model, which readers use to monitor their comprehension (Zwaan et al., 1995; Zwaan & Radvansky, 1998).

Because a coherent situation model requires an integration of what the text explicitly states and the prior knowledge of the reader, it is safe to say that the situation model exists where L1 and L2 reading overlap. This overlap between skills in the L1 and those in the L2 is known as the common underlying proficiency (Cummins, 2000). In this study, I utilized Kintsch's situation model theory to explore the necessary conditions for L2 learners to create a situation model while reading. Because situation models contain an index/slot as part of the text representation that readers use to keep track of the causes of events and the protagonist's intention(s)/goal(s), I also employed Zwaan and Radvansky's event index model. In addition, I incorporated the idea of linguistic threshold (i.e., L2 readers need sufficient L2 proficiency in order to use their L1 reading skills) proposed and argued by Clarke (1980), Alderson (1984), and Carrell (1991) to examine the extent to which L2 proficiency predicts reading comprehension. Theoretical models by Kintsch, Zwaan and Radvansky in conjunction with Clarke and colleagues' linguistic threshold hypothesis allowed me to frame my research questions concerning how educators can target situation model-level comprehension in L2 with

educational intervention.

Statement of the Problem

The situational model, or mental model of discourse is what we bring into our long-term memory (LTM). Cognitively speaking, learning occurs when students make connections between new information and their prior knowledge. Normally, the information used in working memory (WM) is not retained after reading unless it is integrated and stored in LTM. Because no amount of writing can describe every single detail in a situation, readers need to draw inferences by using their LTM to fill in the gaps. Grebe (2009) proposed inferences as a means to process new information to “create a coherent interpretation of academic texts” (p. 70). The question then becomes a matter of how we can help L2 readers integrate text information with their prior knowledge and skills that they possess in L1. Whether L1 inferential ability positively transfers to L2 has not been addressed in previous research on reading comprehension in L2.

Several theorists claim L1 skills transfer to L2 in general. Krashen (1981) views L1 influence as a strategy that learners can build on to acquire L2. Cummins (2000) has expanded this theory to include the idea that L1 literacy skills transfer to literacy in L2 to form the foundation of ELL students’ academic success. Hakuta’s (1990) transfer study showed global transfer of skills from L1 to L2 (as opposed to transfer of specific grammatical forms); students with highly developed overall language skills in Spanish were better able to develop their ability in English. Greater comprehension skills in L1 also transfer to processing information in L2 reading regardless of the linguistic differences between the two languages (Goldenberg, 2008). However, very few researchers apply the concept of situation models to the commonality between L1 and L2 reading. Thus, it is reasonable to apply earlier findings on situation models to observing how ELL students develop their abilities to draw inferences while reading. The same

can also be applied to see how such a process differs between L2 readers with different proficiency levels (e.g., intermediate and advanced readers) in readers' L2.

Purpose of the Study/Research Questions

The purpose of this study was to extend research on situation model-level comprehension to L2 literacy, which had not yet been well studied. This made the study both theoretically interesting as well as highly applicable to instruction on the concept of situation models during reading. The following research questions guided this proposed study: (1) To what extent is ELL students' ability to make L2 inferences affected by their L2 proficiency (Interaction between question type and L2 proficiency), (2) Can engaging ELL students in explaining target situation models (e.g., why and how a certain event occurred in the text as well as what the event caused) improve their L2 comprehension? (Main effect of intervention), and (3) Does the effect of intervention depend on L2 proficiency? (Interaction between intervention and L2 proficiency)

Significance of the Study

The significance of this study includes both theoretical and practical aspects. If the effect of L2 proficiency on inferencing ability is a major factor in determining ELL students' ability to create coherent situation models, lower proficiency ELL students will rely mostly on non-inferential information to comprehend the text in English. In contrast, higher proficiency level ELL students will make not only non-inferential information, but also high-level inferences (e.g., pragmatic inferences). This would confirm that pragmatic inferences occur at a situation model level, while also providing evidence for the linguistic threshold hypothesis, which is discussed in Chapter Two.

Regarding the effect of the intervention, if questioning and explaining makes a difference in creating situation models, the findings from this study would have practical significance. ELL

students in the experimental group would be more likely than those in the control group to comprehend at a situation model level regardless of their L2 proficiency. However, student performance might well be affected by the interaction of both factors (i.e., L2 proficiency and the result of questioning/explaining). Thus, the results of the study would clarify the relative importance for L2 reading comprehension of increasing L2 proficiency in general, versus training on how to make appropriate inferences through questioning/explaining, or the combination of both. New knowledge such as this would benefit a wide range of professionals, including teacher educators, curriculum coordinators, and teachers in terms of how to approach their literacy instruction. Ultimately, such knowledge would benefit their students, whose comprehension might be increased through questioning and explaining at a situation model level.

Limitations of the Study

Due to the fact that participants could not be randomly selected into a classroom within their curriculum, this study employed a variation of an experimental research design. Although having both qualitative and quantitative measures with such designs increased the rigor of the study, limitations of this study included (1) participant selection and (2) use of standardized testing to determine the proficiency level of the participants in their L2. When selecting participants, there is always a chance that groups selected may be different prior to any intervention, which introduces threats to internal validity (i.e., whether an experimental treatment/conditions makes a difference or not, and whether there is sufficient evidence to support the claim). A much stronger study can result when researchers combine both elements (i.e., the comparison between a control group and an experimental group and the evaluation of the pretest-posttest performance on both control and experimental groups, which is known as the

pretest-posttest control group design). Due to its requirement for the random assignment of two groups who receive different treatments, this design makes a true experimental design (as opposed to a quasi-experimental design). It controls various factors that threaten internal validity in the previously mentioned designs, such as history, maturation, testing, and mortality. That being said, experimental research is not easy to conduct in educational settings because of the difficulty in random assignment. Without randomization, we cannot assure that “any systematic variation between experimental conditions is due to the manipulation of the independent variable” (Field, 2009, p. 17).

By using standardized test scores such as EPT (English Proficiency Test) scores to determine the proficiency level of the participants, I was able to compare across proficiency levels how well participants were able to draw inferences based on situation models. Reliability (i.e., no matter how many times being tested, the students’ scores do not vary) is established in such commercially available testing that is widely administered. However, the downside of relying on standardized testing, which is typically done via paper and pencil multiple-choice tests, is that it may not measure a complete picture of one’s linguistic ability

The subsequent section of this chapter lists definitions of key terms concerning text and discourse comprehension that are critical to the literature of this field as well as to this study.

Definition of Terms

Adjunct questions - “questions added to instructional text” (Kintsch, 2005)

Anaphoric reference - a type of reference cohesion in which some current expression refers “back to something previously mentioned in discourse” (Carroll, 2008, p. 161)

Bottom-up processing - makes use of information at the lowest level (e.g., phonemes, letters, words) to perform the highest level of processing

Bridging/connecting inference - a type of inference that establishes a causal relation between two entities (i.e., a focal statement and its antecedent)

Coherence - the degree to which different parts of a text are connected to one another (i.e., Carroll, 2008, p. 423)

Cohesion - “local coherence relationships between adjacent sentences in discourse” (Carroll, 2008, p. 423)

Discourse – “a group of sentences combined in a meaningful manner” (Carroll, 2008, p. 424)

Elaboraters – a type of readers who are likely to draw elaborative inferences from the text

Episodic memory - memory for personally experienced events

Event index model - a comprehension model developed to fully understand “the components and the processes involved in situation model” (McNamara & Magliano, 2009, p. 321)

Linguistic threshold hypothesis - hypothesis proposes that a certain threshold level of L2 language proficiency is necessary before L1 reading ability transfers to L2 reading.

Long term memory (LTM) - memory structure that permanently holds our knowledge, including both episodic memory and semantic memory

Paraphrasers - a type of readers who simply repeat what the text says

Pragmatic/elaborative inference - a type of inference that integrates information from outside of the text, or general knowledge, with that in the text “to fill in missing details and to help formulate a coherent representation of the text” (Oakhill & Cain, 2007, p. 56)

Propositions - idea units containing predicates

Reinstating inference - a type of inference that connects “the focal event to prior text” when antecedent information is not available in the current text (van den Broek, 1994, p. 561)

Semantic memory - memory for generalized facts: “stored information about the features and attributes that define concepts and the processes that allow us to efficiently retrieve, act upon and produce this information in the service of thought and language” Martin & Chao, 2001, p. 194)

Situation model - A mental representation of the situation described by a text

Situation model representations - the mental models of discourse that readers bring into their long-term memory, where prior knowledge is stored

Surface level representations - involve verbatim or exact wording, which can be stored in working memory

Textbase/propositional representations - meaning units represented in the text.

Text-boundedness - readers’ reliance on what is explicitly stated in a text

Think-aloud - a strategy in which the reader focuses his or her attention on the meaning of information by talking aloud while reading; text and discourse researchers use it as a measure to assess representations in the reader’s mind

Top-down processing - making use of information at the highest levels to process language at the lower levels (e.g., guessing the meaning of a word from the context in reading); on the other hand, bottom-up processing

Working memory (WM) - information that is processed and stored while performing cognitive tasks, including reading

Chapter 2 - Review of the Literature

There is no doubt that situation models play a significant role in comprehending texts. Previous research on text and discourse comprehension has shown that readers make use of their general knowledge of the world to create a mental model in order to achieve a coherent understanding of the text that they are reading (Kintsch, 2007; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998; Zwaan & Rapp, 2006). As mentioned previously, this type of mental representation requires the integration of propositional textbase representations (i.e., meaning units represented in the text) and prior knowledge (Brown, 1998; Horiba, 1996; Zwaan & Radvansky, 1998; Zwaan & Rapp, 2006).

In this section of the study, I reviewed previous research on how successful readers, compared to struggling readers, cognitively process information in the text and interact with it. The role of cognitive processing during reading is discussed in relation to readers' general domain knowledge and differences in processing among readers with varied L2 proficiency. I also discussed the role that questioning and explaining can play in the creation of an appropriate situation model during reading, and in the context of an educational intervention that teachers might use to engage their students in coherent discourse. The discussion in this chapter is organized into the following sections: (1) cognitive processing and general domain knowledge, (2) different levels of text representations, (3) top-down and bottom-up processing, (4) inferences generation, (5) the role of working memory in drawing inferences, (6) coherence vs. cohesion, (7) implications of inference drawing abilities for ELL students and their L2 proficiency, and (8) interaction with text, (9) questioning and explaining as a strategy to create situation models, and (10) knowledge base vs. language proficiency.

Cognitive Processing and General Domain Knowledge

What makes readers successful in comprehending text? In their narrative discourse analysis, Trabasso and Magliano (1996) argued that readers achieve coherence through global processing when they are able to integrate “story sentences in a connected memory representation” (p. 161). The researchers assessed school-aged children’s reading comprehension using the think-aloud method, which focused child readers’ attention on each sentence, what it referred to, and how it related to other sentences. Their findings revealed that the method promoted not only young readers’ comprehension but also their monitoring of their own understanding of the text, leading to their holistic interpretation of what they were reading.

Along with the think-aloud protocols, Trabasso and Magliano (1996) were able to assess the readers’ ability to explain and predict what the text was about by asking why and how questions in context. For example, when asked to explain why/how a certain event occurred in the story, the readers were able to retrieve their memory of the prior text and the thought they had associated with the text. This causal relationship between questioning and memory representations during reading has been confirmed by other researchers (Cain & Oakhill, 2007; Graesser, McMahan, & Johnson, 1994; Oakhill & Cain, 2007; Shimizu, 2005). Thus, by having students attend to their cognitive processing (i.e., verbalizing what they understand about the text through making connections between old and new information and drawing inferences based on their situation models by answering why/how questions), the teacher can help them integrate the information stated in the text with the mental representation stored in their long-term memory.

Good readers make use of adequate general domain knowledge to improve comprehension. Readers with rich knowledge structures are more likely to predict and infer what the text is about than those with limited knowledge structures (Afflerback, 1990; Garcia, Jimenez, & Pearson, 1998; Grabe, 2009; Graesser, Singer, & Trabasso, 1994; McNamara &

Kintsch, 1996). Graesser and colleagues (1994) argued that such knowledge structures are grounded in readers' experiential knowledge (e.g., having restaurant schema/script). A schema refers to "a structure in semantic memory that specifies the general or expected arrangement of a body of information" (Carroll, 2008, p. 176). For example, one's familiarity with the topic or content affects comprehension in terms of its automatic activation in working memory. Consider the following passage:

Donald and his third wife entered Mario's. When they were seated, Donald declared himself unhappy with his table, which was located near the door, was not conducive to a romantic conversation. He requested a table in the back of the restaurant. He studied the menu and ordered ossobucco for two. The food was great. Instead of a dessert, they ordered cappuccino. Donald left a big tip. (Zwaan & Rapp, 2006, p. 733)

The passage does not explicitly say that Donald and his wife are at the restaurant, but most readers have no trouble understanding this because they can rely on their restaurant schema (some researchers referred to it as "script"). On the other hand, readers who have never dined at a restaurant may not be able to easily process information such as being seated, ordering from the menu, and tipping. Teachers who serve students need to consider supplying students with information they may be missing (e.g., who seats customers, takes orders, delivers food, receives the tip).

How information is presented also makes a difference in reading comprehension. The following passage might be difficult to process for those who are not provided with its title, "Christopher Columbus Discovering America":

With hocked gems financing him, our hero bravely defied all scornful laughter that tried to prevent his scheme. "Your eyes deceive." He had said, "an egg not a table correctly

typifies this unexplored planet.” Now three sturdy sisters sought proof, forging along sometimes through calm vastness, yet more often over turbulent peaks and valleys. Days became weeks as many doubters spread fearful rumors about the edge. At last from nowhere welcome winged creatures appeared signifying momentous success. (Dooling & Lachman, 1971, as cited in Carroll, 2008, p. 177)

This example illustrates the importance of a text being presented with its title as well as the necessity of schema development about the topic (e.g., who Christopher Columbus is). A study by McNamara and Kintsch (1996) revealed that readers who were familiar with text content (e.g., Vietnam War) performed better on the essay questions after reading low coherence texts, such as the text below, than those who lacked such domain knowledge.

By the fall of 1964, Americans in both Saigon and Washington had begun to focus on Hanoi as the source of the continuing problem in the south. As frustrations mounted over the inability of the ARVN to defeat the enemy in the field, pressure to strike directly at North Vietnam began to build. (Britton & Gulgoz, 1991, as cited in McNamara & Kintsch, 1996, p. 255)

Classroom teachers may want to have their students brainstorm about their knowledge of the historic figure or the event (i.e., discovery of America) through mind mapping (i.e., a way of organizing ideas) prior to and during reading to enhance comprehension. The more readers make connections in such a way, the more information is retained as semantic memory. Similarly, Grabe (2009) suggested “using semantic mapping and making explicit connections between background knowledge and specific details of the text” as a way to connect information in academic texts with students’ domain knowledge (p. 82).

Different Levels of Text Representations

What does it mean to comprehend a text? Various cognitive psychologists and discourse researchers identify three basic levels of memory representations that readers need to construct during reading: surface representations, textbase/propositional representations, and situation model representations (Brown, 1988; Carroll, 2008; Kintsch, 2007; Rai et al., 2011; Shimizu, 2005; Zwaan & Radvansky, 1998; Zwaan & Rapp, 2006; Zwaan & Singer, 2003). Surface representations involve verbatim or exact wording, which can be stored in working memory (i.e., information that is processed and temporarily stored while performing cognitive tasks, including reading) (Carroll, 2008). For example, readers are better able to recall the exact information contained in the current sentence than in a previous sentence. This is particularly true for L2 readers, as working memory is a limited resource in their L2 (Schwartz & Kroll, 2006).

Textbase/propositional representations refer to the meanings of sentences, which are stored apart from the exact wording but in the form of propositions or meaning units. The more propositions a sentence has, the longer it takes to read regardless of the number of words contained in the sentence. Compare the following example sentences (Carroll, 2008, p. 169):

“Cleopatra’s downfall lay in her foolish trust in the fickle political figures of the Roman world.”

“Romulus, the legendary founder of Rome, took the women of the Sabine by force.”

These sentences contain nearly the same number of words, but due to the greater number of propositions, the first sentence is far more complex and difficult for readers to process than the second sentence. In short, the more propositions a sentence has, the more working memory is required for comprehension. Although it is necessary for readers to construct propositions from a text, current researchers (Brown, 1998; Foltz, 2003; Graesser, Singer, & Trabasso, 1994; Horiba, 1996; van den Broek, 1994; van Dijk & Kintsch, 1983; Zwaan, Langston, & Graesser, 1995;

Zwaan & Radvansky, 1998, Zwaan & Rapp, 2006; Zwaan & Singer, 2003) have found the meanings represented in such a textbase level rather limited for achieving high levels of discourse comprehension. Global mental processing (i.e., the situation model) is needed for the reader to fill gaps with background knowledge in order to fully comprehend.

Situation model representations are the mental models of discourse that readers bring into their long-term memory, where prior knowledge is stored. According to van Dijk and Kintsch (1983), there is a strong connection between a situation model and long-term memory, or a memory structure that permanently holds our knowledge, including both episodic memory (i.e., memory for personally experienced events) and semantic memory (i.e., memory for generalized facts).

There are strong theoretical reasons why memory should be better for the situation model than for the textbase itself: Retrieval is most likely to occur if a memory episode is embedded within some larger structure which can serve as a retrieval system. Situation models, by their very nature, tend to be embedded in such systems and form a part of a larger model. (van Dijk & Kintsch, 1983, pp. 340-341)

Consider the following passage to see how readers must create and recreate a mental representation of the situation in order to comprehend each sentence:

Peter took the elevator to the fifth floor. He went to talk to his professor. He was anxious to find out how the professor liked his draft. He walked up to the professor's office and knocked on the door. The professor looked up from his work. (Zwaan & Radvansky, 1998, p. 166).

In sentence one, readers start out by creating a situation model in which Peter rides an elevator for some unknown reason, assuming that he is in a building. As readers move to the second

sentence, they understand the purpose of Peter's riding the elevator (i.e., visiting his professor), while inferring that Peter is a student. By the time readers complete the third sentence, they capture a better picture of the situation, including Peter's emotional state. The fourth and fifth sentences provide readers with a causal relationship between Peter's action of knocking and the professor's action of looking up. Thus, readers not only construct the initial situation model, but also continue to update it by relying on both textbase propositions and their general background knowledge (i.e., elevators being in buildings, knocking on the door causing a response if someone is behind the door). The event in the above passage is indexed by time (during office hours), space (university building), causality (Peter's knocking and his professor's looking up), intentionality (Peter's wanting to talk to his professor about his paper), and protagonist (Peter). In addition, general knowledge of the world is activated and integrated with the text, which allows them to continuously make connections among different narrative dimensions to construct a coherent understanding of the passage.

When reading a text, readers continually update their comprehension according to each situation that they encounter. This is where the unfixed nature of situation models is especially beneficial. Because such mental representations are dynamic and ever changing as the reader continues to read, pedagogical activities during reading should occur simultaneously and allow readers to revise their situation models to construct a coherent understanding of the text. According to the event-indexing model proposed by Zwaan et al. (1995) and Zwaan and Radvansky (1998), readers keep track of narrative dimensions such as time, space, cause/effect relationships, protagonist, and the protagonist's goal. Consider the following example:

George got a peanut butter jar out of the cupboard and put it on the kitchen counter. He opened the jar and put the content on a butter knife. He carefully put a small amount of the peanut butter on a tip of a mousetrap.

Until readers come to the final sentence, they are likely to have created the situation model of George (protagonist) making a peanut butter sandwich (goal) in the kitchen (space). Upon reading the final sentence, they immediately need to revise their understanding of George's goal to that of setting a mousetrap. As the above example illustrates, mapping new information onto existing information during reading requires that knowledge structures be coherent. When they become incoherent, readers need to build new situation models (e.g., setting a mousetrap). Because situation models are constructed at the time readers encounter events in the text, it makes sense to incorporate them into during reading activities.

Brown (1998) described the situation model as “an active mental structure that is the result of an interaction between the reader and the text” (p. 193). In her view, each reader is likely to activate his/her unique prior experiences to make sense of the text, and such a knowledge structure can be updated as the reader receives new information. Foltz (2003) also described this non-fixed nature of knowledge represented in the situation model as “a function of the task and of the person's current activated knowledge” (p. 509). Using a computational linguistic technique called “latent semantic analysis” (LSA), Foltz and others (Foltz, 2003; Foltz, Kintsch, & Landauer, 1998; Grasser et al., 2000; Magliano & Millis, 2003) have measured how readers make semantic associations among words and sentences they are reading—representative of their current situation model—to better comprehend texts.

Top-down and Bottom-up Processing

Top-down and bottom-up processing are two different ways we process language. The

Lamar Alexander example from Chapter One illustrates how top-down activation of world knowledge enables a reader to draw pragmatic inferences that are needed to form a coherent situational model and thus engage in active comprehension (Zwaan & Radvansky, 1998). From an educational point of view, Smith (1988) defined top-down processing as a strategy in which “the reader determines how a text will be approached, dealt with, and interpreted,” as opposed to its passive counterpart, bottom-up processing, which involves “putting the text in charge, with the letters on the page first and final arbiters of the reader’s response” (p. 218). According to Carroll (2008), top-down processing refers to making use of information at the highest levels to process language at the lower levels (e.g., guessing the meaning of a word from the context in reading); on the other hand, bottom-up processing makes use of information at the lowest level (e.g., phonemes, letters, words) to perform the highest level of processing.

ELL students often have weaker mental representations in English than native English-speaking students due to their rather limited abilities to use linguistic cues to construct appropriate situation models. Therefore, L2 readers should consciously use top-down processes as a strategic approach to creating a situation model to comprehend text, especially with unknown words and grammatical structures. Horiba, van den Broek, and Fletcher (1993) analyzed how L2 readers filled in gaps in their mental representation after reading narrative texts. The researchers found that the readers relied on top-down, filling-in processing (i.e., preserving structural properties of the texts such as causal relationships and story-grammar) to recall and approximate the meaning of unknown words in the text rather than bottom-up processing (i.e., the constituent elements of the text). This finding suggests that when their L2 is weak and their comprehension breaks down due to a lack of lexical knowledge, L2 readers are more likely to make use of a more holistic process at their lower proficiency level, which

facilitates construction of the meaning represented in recall. “Because students have few other resources, the situation model provides an opportunity to impose some degree of coherence” (Grabe, 2009, p. 49).

As mentioned earlier, inferences are necessary in language comprehension in order to retain coherence because information is always ambiguous, especially for L2 readers. Among various types of inferences, ranging from bridging to elaborative inferences, van den Broek (1994, p. 574) argued that those which “provide coherence to the text by connecting the focal statements to prior events” are the most crucial to resolving anaphoric ambiguity (i.e., difficulty in identifying what each expression is referring back to prior information).

By having her participants report their thoughts in their native language, or L1, during the reading process using a measure, think-aloud, Horiba (1996) found that L2 readers used different inference generation processes (i.e., reinstating and connecting/bridging, reinstating, and elaborative inferences) depending upon their proficiency level in their L2. A connecting/bridging inference establishes a causal relation between two entities (i.e., a focal statement and its antecedent), whereas a reinstating inference refers to connecting “the focal event to prior text” when antecedent information is not available in the current text (van den Broek, 1994, p. 561). Graesser et al. (1994) described the reactivation involved in reinstating inferences noting that “reinstatement searches are executed to fetch the earlier text and to place it in working memory (WM)” (p. 378). Readers use their general background knowledge to draw an elaborative inference in order to make connections between the focal event and an unmentioned antecedent. According to the verbal responses that Horiba (1996) observed, high-proficient L2 proficient readers made text coherence more through reinstatements, which require accessing information from long-term memory (LTM) representations, whereas low-proficient readers did so more

through connecting inferences activated within short-term memory (STM) representation. Her L2 advanced readers “generated backward connections between ideas in the current sentence and the prior text more effectively” than her intermediate readers (Horiba, 1996, p. 460). The researcher concluded that both levels of text representations (i.e., propositional textbase representation and situational model) are necessary to achieve coherence in L2 reading.

In the case of L2 reading, research shows conflicting evidence of readers’ overreliance on top-down processing to make sense of the meaning of the text (Horiba, van den Broek, & Fletcher, 1993; Johnson, 1981; Steffensen, Joag-dev, & Anderson, 1979) and on bottom-up processing due to their inability to utilize the knowledge structures effectively (Carrell, 1983, 1991; Nassaji, 2007). Carrell (1984) investigated underlying causes of such contradiction in L2 reading, finding that variation is likely due to schema availability, schema activation, and skill deficiencies. She found that L2 readers were more likely to stay at the textbase level due to (1) a lack of appropriate background knowledge, (2) an inability to activate background knowledge even if they have it because the text itself doesn’t provide readers with cues to use background knowledge, and/or (3) a lack of linguistic knowledge. Thus, text-boundedness (i.e., readers’ reliance on what is explicitly stated in a text) can occur for reasons related to both top-down processing (as in lack of background knowledge) and bottom-up processing (as in lack of linguistic knowledge). Clearly, Carrell’s findings support the linguistic threshold hypothesis (i.e., learners need to read at a threshold level of L2 proficiency before being able to transfer skills and knowledge that they possess in their L1). How might this theory apply to skills such as generating inferences while reading?

Inference Generation

Various researchers (e.g., Afflerback, 1990; Garcia, Jimenez, & Pearson, 1998; Grabe 2009) have proposed teaching inference-making skills as a strategy for L2 readers to better comprehend text. After observing how high knowledge readers made use of inferences to construct the main idea of a text and to monitor their comprehension, Afflerback (1990) regarded activating prior knowledge as an important factor in drawing accurate inferences. In their study comparing successful monolingual and bilingual readers, Garcia et al., (1998) found that both groups of children made use of checking the inferences that they made with textual information in order to achieve comprehension, while also constantly revising their comprehension. Observe the way each of the bilingual children drew a higher-level inference about the outcome/information of the story in the following utterances:

Alberto: There might be a chance that I finish the story and ... all human beings might be extinct for all the pollution and stuff.

Glida: Oh! Now I know what he's going to do! So the coat is not his; it belongs to the man. (Garcia et al., 1998, p. 207)

Less successful bilingual readers, on the other hand, forced “subsequent text information to fit earlier interpretations” without making appropriate inferences or revising their comprehension.

The following example illustrates this problem:

Celina: Well, it said it's a man and I don't think it was a man cuz a man couldn't be more dangerous than an elephant or a tiger or a bear (Garcia et al., 1998, p. 207).

Celina's thinking process did not go beyond what was explicitly stated in the text.

Zwaan and Rapp (2006) considered how learners' background knowledge is organized in their long-term memory through semantic connections. The researchers proposed that readers' knowledge is retrieved by inference generation during their reading comprehension. Essentially,

there are two views of inference processing: (1) strategic processing and (2) automatic activation. The former views comprehension as an active process in which readers consciously generate inferences (Graesser, Singer, & Trabasso, 1994), whereas the latter views it as a result of passive activation of knowledge structures (McKoon & Ratcliff, 1992; Myers & O'Brien, 1998). Zwaan and Rapp (2006) argued that, alone, neither view is completely supported by research and that regardless of the view one incorporates, comprehension is always incomplete. Therefore, both views of inference processing could take place at the situation model level, with the only difference being conscious/subconscious effort.

As mentioned earlier, memory representation is better at the situation model level because it involves readers' tracking of time, space, cause/effect relationships, protagonist, and the protagonist's goal (Zwaan, et al., 1995; Zwaan & Radvansky, 1998). By integrating textual information, background knowledge, and cognitive activities (e.g., drawing inferences and keeping track of each event with their memory representation), readers can achieve comprehension (Grabe, 2009; Zwaan & Rapp, 2006). An earlier study by Hansen (1981) and more recent longitudinal studies by Oakhill and colleagues (Cain & Oakhill, 2007; Oakhill & Cain, 2007; Oakhill & Yuill, 1997) have shown the significant role of inference-making tasks and prior knowledge connections in readers' comprehension.

The Role of Working Memory in Drawing Inferences

Needless to say, drawing inferences requires a sufficient amount of working memory (WM), or storage for retaining information. For example, Singer (1994) argued that the higher WM one has, the more bridging inferences take place to process antecedent information. Bridging inferences, including anaphoric reference, cataphoric reference, and pronominal reference, allow cohesion and reflect the role of WM in discourse comprehension. For example,

anaphoric reference refers to a type of reference cohesion in which some current expression refers “back to something previously mentioned in discourse” (Carroll, 2008, p. 161). Observe the following sentence: “I did my homework. The algebra was very hard.” “The algebra” in the second sentence is an anaphor referring to the homework in the previous discourse (i.e., the first sentence.) Cataphoric reference uses referring expressions to refer forward as in “This is what you want to do. First, write a topic sentence followed by supporting details.” “This” in the first sentence is a cataphor, which refers to the information that follows (Carroll, 2008). Pronominal reference, as in “Sara left me a message; she said to call back” means that the pronoun “she” refers back to “Sara.” All the above examples are concerned with connecting information within the same text, and research has shown that readers with a low working memory span predominantly rely on bridging inferences to make sense of words used in different parts of discourse (St. George, Mannes, & Hoffman, 1997; Rai et al., 2011).

Elaborative inferences are another type of inference used by readers to interpret ambiguous discourse. By adding information, readers can better elaborate on their situation model of the text they are reading. However, Whitney and colleagues’ study (1991) indicated that readers with a low working memory span immediately draw more specific elaborative inferences, whereas readers with a high working memory span draw more general inferences, leaving their interpretation open ended until the end of a sentence/paragraph.

WM is limited for all readers regardless of their L1 “because reading is accomplished too quickly for some time-consuming inferences to be generated” (Graesser, Millis, & Zwaan et al., 1997). However, WM is a particularly limited resource for L2 readers due to the linguistic gaps between their L1 and L2 (e.g., vocabulary size; exposure to discourse; implicit knowledge in syntax, morphology, and phonology). For example, English language

learners cannot simply deduce the meanings of phrasal verbs by studying each constituent, and this makes it hard for them to process the information contained in the phrasal verbs. In addition to its formulaic and idiomatic nature, the constituents of a phrasal verb can be separated depending upon the particular object of the verb, which contributes to further difficulty in processing (Folse, 2009). To comprehend sentences such as “The coach called all seven of the season’s games with tough Division Three schools off,” readers need to understand that “called” and “off” are parts of the same unit that are separated by the entire noun phrase “all seven of the season’s games with tough Division Three schools,” which is the object of the sentence. Due to the distance between the two words, readers need to go back to the beginning of the sentence to make sense of the sentence, while their working memory lasts. This processing issue seems to parallel the effect that a long distance between an antecedent and its anaphor has on working memory in comprehension, as in Carroll’s example of bridging inference:

“Yesterday a black dog bit a little girl (antecedent). It got away, and we are still trying to find it. He is short and has a dog tag on his neck that says Fred. She (anaphor) was scared, but she wasn’t really hurt.” (2008, p. 164).

Van den Broek and Espin (2012) framed the causal relationship between WM and inferencing within their “three clusters of higher-order cognitive factors and processes”. The framework consisted of general cognitive factors and processes (including WM capacity), comprehension factors and processes (including inference generation), and text comprehension factors and processes. The researchers used the integrated model to describe the complexity of reading comprehension processes in which clusters of cognitive, comprehension, and text

comprehension factors are connected with one another. For example, WM allows readers to monitor their comprehension, while making inferences and building coherence. Then, they strategically incorporate their specific knowledge of text structures such as narratives and expository writing into guiding their overall reading comprehension.

Coherence vs. Cohesion

Working memory also plays a major role in another type of discourse inference: pragmatic inference. Let us revisit the earlier example of pragmatic inferencing (e.g., “John dropped a vase; therefore, he spent a lot of money to replace it.”), which has to do more with semantics or how we infer meaning (John broke the vase). Although the initial sentence does not explicitly state the inference, the writer implicitly communicates the intended meaning to the reader. Unlike bridging inferences, the reader needs to integrate information from outside of the text, or general knowledge, with that in the text “to fill in missing details and to help formulate a coherent representation of the text as a whole” (Oakhill & Cain, 2007, p. 56). According to Graesser et al. (1994), such inferences are used to maintain coherence (i.e., “the degree to which different parts of a text are connected to one another”) rather than cohesion (i.e., “local coherence relationships between adjacent sentences in discourse”) (Carroll, 2008, p. 423).

Although the two concepts—coherence and cohesion—are related, they are independent from each other. Consider the following cohesive passage without coherence:

In a little Danish town, two fishmongers exchange blows. Anders, by far the stronger had a cousin in prison. Anders was twice the age of the cousin. When he first was convicted, Anders was living in Italy. Anders has a wife who lost her bathing cap. Her car is at this moment double-parked. (Samet & Schank, 1984 in Zwaan & Radvansky, 1998, p. 164)

What is wrong with this passage? Although each sentence is connected to its preceding sentence through reference cohesion (e.g., Anders equals one of the fishmongers, the car doubled-parked was Anders's wife's), there are no coherent connections between the sentences that readers can draw on to comprehend the whole passage as a meaningful unit. Thus, there are no integrated situation models from which they can draw inferences. This suggests that readers need to go beyond just making low-level inferences, such as an anaphoric reference (i.e., a type of reference cohesion in which some current expression refers back to that of something previously mentioned). Rather, they need to construct a coherent representation of the story (high-level inferences) to achieve reading comprehension.

Implications of Inference Drawing Abilities for ELL Students

The likelihood of drawing such low-level inferences or bridging inferences among readers with less L2 proficiency correlates with their lower language competence or lower WM span (Horiba, 1996). Horiba's findings were consistent with other second language acquisition (SLA) studies on inferences (e.g., Hammadou, 1991; Kembo, 2001; Rai et al., 2011; Shimizu, 2005). Hammadou (1991) looked at the relationship among three variables: readers' prior knowledge, their ability to draw inferences, and their language proficiency in L2. She found that there was no difference in the amount of self-reported prior knowledge activated between L1 and L2 readers, but compared to their L2 counterparts, L1 readers tended to recall more non-textual propositions, which required generating appropriate inferences from their general knowledge. Thus, this study revealed that the more proficient readers were in their target language, the more successful they were in logically making connections between the text and their knowledge of the world. In other words, there is an effect of language competence on an inference process, which involves sufficient reasoning through generalization. Gernsbacher (1993) and Gernsbacher

and Robertson (1995) also confirmed that skilled readers are able to identify and reject inappropriate information in order to comprehend texts better and faster than less skilled readers. Their findings mirrored the conclusion that proficient readers make more appropriate inferences than less proficient readers in L2 (Long & Golding, 1993; Whitney, Richie & Clark, 1991).

Kembo's (2001) study revealed that child L2 readers had difficulties drawing pragmatic inferences, which required a greater level of text integration compared to other types of inferences that were textually explicit (e.g., pronominal and logical informational inferences). For example, the students were able to respond to the questions "What does her refer to?" and "What should Maria have done?" after reading the text "Maria was driving without lights when the police stopped her" more easily than questions that required elaborative/pragmatic inferences such as "What kind of driver is Maria?" This type of question demands that readers come up with an answer after integrating the different propositions within the text, while incorporating their general knowledge of the world. These results were consistent with the readers' own judgment of question difficulty according to the one-to-five scale rating (i.e., very easy, easy, between easy and difficult, difficult, and very difficult). The majority of the L2 readers found pragmatic inference questions to be the most difficult to process.

The research of Rai and colleagues (2011) on adult foreign language learners showed that low-WM-span readers predominantly rely on bridging inferences to make sense of words used in different parts of discourse. The researchers investigated how stress impedes foreign language comprehension due to interactions between readers' WM capacity and types of inferences that were required to comprehend the text. The purpose of the study was to explain how stress impedes the cognitive process of foreign language reading in relation to the role of WM and its interaction with inferential complexity. They predicted that stress would impede their

participants' reading comprehension with stronger effects on pragmatic inferences, which require a higher level of text representation (a situation model) than bridging inferences and non-inferences (which use only explicitly stated information provided in the text). The researchers' multiple comparisons showed that there was "significantly greater accuracy for non-inference ... questions than for either bridging ... or pragmatic inferences, and a marginally significant trend for more accuracy for bridging inferences than for pragmatic inferences" (Rai et al., 2011, p. 25). In short, a higher level of inferential skill, which requires greater WM makes reading processing more difficult for foreign language learners.

Interaction with Text

Discourse comprehension studies have found that active interaction between texts and readers facilitates greater comprehension (Garcia et al., 1998; Graesser, Singer, & Trabasso, 1994; McNamara, 2004). The ability to relate new information in the text to the information that readers already have in their long-term memory seems to be a defining characteristic of what makes a good reader. Garcia and colleagues (1998) observed how good readers constantly monitored and revised their current comprehension whereas struggling readers did not. Less successful readers in the study were stuck at the textual information level and did not make connections to prior knowledge or revise their comprehension. In addition to monitoring comprehension, McNamara (2004) found that students who used a self-explanation strategy (i.e., explaining aloud the meaning of information to themselves while reading) showed greater reading comprehension than those who simply read aloud. McNamara observed the effectiveness of the strategy, especially among students with low domain-general knowledge. This finding indicated that the process of explaining textual meaning supports readers in filling knowledge gaps, resulting in increased comprehension. This strategy is particularly important when students

read expository texts, which are more likely to include not only unfamiliar information but also less cohesive information (McNamara, 2004).

The ability to make decisions about whether information is relevant to the key point of a text also separates skilled readers from less skilled readers. In their studies, Gernsbacher (1993) and Gernsbacher and Robinson (1995) found that skilled readers were more likely to suppress irrelevant information in order to build a coherent mental structure, which the researchers refer to as a “structure building framework,” for better comprehension. Suppression as a strategic reading process was also evident in work by Long, Johns, and Morris (2006). These researchers attributed to working memory capacity the failure of less successful readers to inhibit irrelevant information. Teachers can apply this theory of suppression to guide their students to focus on key information by eliminating unnecessary details while reading.

Questioning and Explaining as a Strategy to Create Situation Models during Reading

Teaching methods that are directly related to the situation model theory of reading comprehension include questioning (Ajudeh, 2003; Carrell et al, 1998; McNeil, 2010) in addition to self-explaining (McNamara, 2004). By having students create their own questions using cue cards containing *who*, *what*, *when*, *where*, *how*, and *why*, the teacher can support students’ ability to attend to the key information needed to increase their comprehension. The teacher might also ask students to write or draw examples of who the protagonist is, what his/her goal is, when and where the action took place, how the goal was achieved, and why the action caused an outcome. However, teachers need to keep in mind that not all questions during reading lead to deeper levels of comprehension. Graesser (2007) and Magliano, Trabasso, and Graesser (1999) pointed out that *how* questions and *what happens next* questions seem to disrupt readers’ processing and

memory. However, to constructionists, explanations and why-questions are fundamental to the construction of meaning (Graesser, 2007).

Shimizu (2005) investigated the inference generation process of adult L2 learners and the effects of reading questions on their inference making. Like Horiba (1996), the researcher incorporated think-aloud tasks, but she also added an oral questionnaire to investigate how L2 learners respond to inference questions (e.g., “Why do you think the event in the last sentence occurred?”). This is consistent with earlier findings about comprehension being guided by why-questions eliciting inferences (Hansen, 1981). In her study with child L1 readers, Hansen asked both inferential and non-inferential questions after having children read the following text:

Sheep dogs must learn to take sheep to eating places. They take the sheep from place to place.

Non-inferential question: Where do the sheep dogs take the sheep?

Inferential question: Why do they take the sheep from place to place? (p. 404)

The why-question allowed the students to draw inferences such as, “because the sheep eat so much grass that the pastures soon become barren” after making connections to their prior knowledge.

According to Graesser, McMahan, and Johnson (1994), why-questions activate “information sources in working memory” (p. 528). In their work on children’s problems with reading comprehension, Cain and Oakhill (2007) and Oakhill and Cain (2007) emphasized that, based on the causal relationship between inference making abilities and children’s reading comprehension abilities, students should practice inference tasks to improve their comprehension skills. The researchers’ longitudinal study showed that inference training such as strategically using wh-questions (though not as limiting as why-questions) and paying attention to clue words

to generate students' own questions increased students' comprehension skills over time.

Graesser, Singer, & Trabasso (1994), however, narrowed the effectiveness of why-questions to specific inference types, such as causal antecedents.

By engaging participants in speaking aloud their thoughts about a passage, answering why-questions, and rereading the same passage, Shimizu (2005) sought to capture the participants' thought processes while reading. The study's findings included: (a) inferences were significantly generated by the questions, which made readers infer the possible causal relationship between particular events within a text, and (b) the questions also allowed the learners to comprehend propositional meanings/idea units of a text. Therefore, Shimizu concluded that in addition to lower-level processing questions (e.g., bridging inferences), higher-level processing questions that require situation models should be included in reading comprehension assessments (e.g., TOEFL).

Magliano, Trabasso, and Graesser (1999) looked at how readers used explanatory and predictive inferences as strategies to integrate information into coherent memory representations and found "a positive relationship between the number of explanations participants generated while reading a text and long-term memory" (p. 626). In other words, participants who were trained to explain the content of the text through questioning had better text retention and thus better comprehended while reading silently. The study pointed to the increase of memory representation as a result of asking why and how questions to support comprehension of the text.

As earlier research suggested, a coherent understanding of the text content leads to deep level understanding, which cannot occur without situation models. E. Kintsch (2005) reviewed multiple roles of questioning in forming situation models, including helping students build their knowledge and assess their learning. The author found adjunct questions (i.e., "questions added

to instructional text”) to be particularly beneficial for learning when they occur shortly after the relevant passage or at logical break points in text (p. 52). Inference questions, which require learners to read beyond what is explicitly stated in the text, should be placed “close to the statement in the text that cues the inference” (p. 60).

Knowledge Base vs. Language Proficiency

The situation model theory assumes that comprehension requires the integration of textbase representations and activated prior knowledge, which is constantly updated by the ongoing situation. To achieve comprehension, readers need to take what the text literally says and connect it to their knowledge base, while constructing meanings through a situation model. How does the situation model theory apply to L2 reading? Regardless of linguistic differences between the L1 and the L2, native speakers and non-native speakers of the target language use the same semantics that are generated by their situation models.

Another set of noteworthy findings is that while earlier research demonstrated the effect on recall of having schematic knowledge (Anderson, Reynolds, Schallert, & Goetz, 1977; Bransford & Johnson, 1972; Carrell, 1983), more recent research has indicated that students must have at least intermediate language proficiency (in the language of the text) to be able to use appropriate knowledge of the topic to improve reading comprehension (Barry & Lazarte, 1995, 1998, 2000; McNeil, 2010). Barry and Lazarte (1995, 1998) analyzed how high-domain-knowledge and low-domain-knowledge readers performed on recall tasks. They found that having domain knowledge provided an advantage for high-knowledge readers, which is not surprising. However, the complexity of sentence structures (i.e., more embedded clauses in a sentence) “seemed to cancel the advantage of previous exposure to the content domain” (Barry & Lazarte, 1995, p. 491). For domain knowledge to have a positive affect on readers’

comprehension, they need to have at least intermediate proficiency in their L2, which enables them to process embedded clauses. If intermediate proficiency enables L2 readers to utilize appropriate schemata, then simply having students activate their prior knowledge is not sufficient to improve their reading comprehension, especially for those at the lower proficiency level. This is consistent with the previously mentioned findings of Carrell (1984), providing evidence for the linguistic threshold hypothesis. Other researchers such as Lee and Schallert (1997), Yamashita (2002), and Uso-Juan (2006), also came to a similar conclusion in their English as a foreign language (EFL) studies: L2 proficiency was the more important predictor for students' successful reading in L2 than their L1 reading ability and/or domain knowledge in their L1.

In addition to the effect of language proficiency and background knowledge on reading comprehension, McNeil (2010) considered another variable—reading comprehension strategies (e.g., self-questioning). In this study involving university-level L2 learners who had completed reading level three out of six of an intensive English program, McNeil found an overwhelmingly strong relationship between use of self-questioning and reading comprehension (this relationship accounted for 56.7% of the variance in reading comprehension scores), but not background knowledge. This study suggests that conscious actions of questioning and monitoring can outweigh the influence of existing domain knowledge (or the lack thereof).

Summary

Due to the complexity of text comprehension, including the fact that it is an ongoing process, readers need to go beyond the textbase level of representation. To enhance use of appropriate situation models among students so that they can comprehend the underlying content of a text, teachers need to provide students with opportunities to interact with text at a deeper level. By utilizing top-down processing, students can rely on their general domain knowledge to

predict what the text is about. While doing so, the students can also rely on bottom–up processing (e.g., lexical information) at the same time. By incorporating cognitive strategies such as questioning and explaining text content, students can construct and reconstruct their situation models, while also monitoring their comprehension to maintain coherency. To assess such abilities, teachers can engage students in generating inferences, especially pragmatic inferences. In the case of L2 learners, these types of strategies can help them to monitor their situation models and compensate for linguistic knowledge gaps and limited working memory in their L2. Given the benefits of situation models in reading comprehension, it is crucial for teachers to train their ELL students to become strategic readers who make use of such resources to construct underlying meaning from L2 text.

As this review of the literature illustrates, very little research has applied the idea of situation models to increasing comprehension in L2 reading. Thus, it is worthwhile to examine the effect of questioning and explaining on creating situation models for L2 learners' reading comprehension. Also needed is investigation of how such processes interact with L2 proficiency. The next chapter describes the methodology of this study, including its research questions and the research design to answer the questions.

Chapter 3 - Methodology

Research Questions

As identified in Chapter One, this proposed study was designed to answer the following research questions: (1) To what extent is ELL students' L2 proficiency a factor contributing to their ability to make appropriate inferences?, (2) Can engaging ELL students in actively explaining target situation models (e.g., why and how a certain event occurred in the text as well as what the event caused) compensate for lack of linguistic competence in the L2?, and (3) Does the effect of intervention depend on L2 proficiency?

Based on these overarching questions, the following hypotheses were proposed:

- H1: Factual/non-inferential questions are more likely to be answered appropriately by L2 learners than those that require the situation model level inferences, which is the highest level of text representation (after surface and propositional representations).

- H2: The more proficient ELL students are in their L2, the better they will make appropriate inferences based on situation models. Similar to Hypothesis 1, the pragmatic inferences are more likely to be made by the learners with higher proficiency than those with lower proficiency, and the factual/non-inferential questions will be correctly answered at equal rates by learners of all levels of L2 proficiency.

- H3: As ELL students actively engage in responding to explanatory/adjunct questions (asking for elaborative information that is integrated at the deeper level), they will better create situation models, and thus better comprehend the text that they are reading.

- H4: The effectiveness of having ELL students respond to explanatory/adjunct questions may differ according to proficiency level. For example, there may be a linguistic threshold below which the explanatory questions will not help. Conversely, if answering explanatory

questions helps overcome L2 proficiency limitations, then it could be that the effect will be greater at an intermediate level of proficiency than at a high level of proficiency.

Research Design

As stated in Chapter One, this study employed an experimental research design to compare data from beginning level ELL students, intermediate level ELL students, and high-intermediate level ELL students as a between-group variable to determine if inference patterns (dependent variable) differed by proficiency level (independent variable). The nature of students' explanations and where the information is coming from (i.e., the evidence that they provide from the text) also was analyzed. The correlation between the quality of explanations (as a categorical dependent variable) and L2 proficiency (as a continuous independent variable) was examined. Such correlations predicted the quality of explanations as performance and outcomes based on proficiency.

Due to the differences in language competence among the three groups, it was natural to predict that low-intermediate L2 readers were less likely to make appropriate inferences than intermediate or high-intermediate L2 readers. Fewer appropriate inferences, in turn, would reduce readers' comprehension. High-intermediate L2 readers were most likely to succeed in generating inferences by making connections between old and new information. However, by engaging students in developing situation models during reading (i.e., having them work through what the event/situation is like via explaining vs. not explaining) as an intervention, the researcher predicted that readers of all proficiency levels in L2 would make more appropriate inferences and thus increase their comprehension.

The researcher asked pedagogical questions to the experimental group to encourage them to explain the underlying situation of the text that they were reading. The control group was

asked to reread the text as an additional task to replace explaining. Thus, this study was a variation of a factorial design (i.e., “any experimental design with more than one independent variable”, Goodwin, 2010, p. 564). In particular, it had a 2 x 2 x 3 design with two question types (non-inferential vs. inferential), two levels of instructional treatment (explanatory questions vs. rereading control condition), and three levels of L2 proficiency (low-intermediate, intermediate, and high-intermediate). The study incorporated a static group comparison involving two non-randomly formed proficiency groups, where one group was exposed to the experimental treatment, and the results were tested while a control group was not exposed to the treatment and was similarly tested in order to compare the effects of treatment (Gay, Millis, & Airasian, 2006).

Population and Sample

One hundred seventeen ELL students attending this state university in northeastern Kansas who happened to be available participated in this proposed study (convenience sampling). All participating ELL students were native speakers of Spanish from Ecuador, which held ELL students’ native language constant and also mirrored the fact that Spanish speakers were the largest group of ELL learners attending this university. The students were admitted to the university to begin college-level work but still needed to meet English proficiency standards for the university (i.e., attain a high enough score on the English Proficiency Test [EPT]). Thus, the majority of the students were enrolled in the English Language Program (ELP), which provided English instruction to meet such linguistic needs of international students.

The study compared data ranging from low-intermediate to high-intermediate level ELL students as a between-group variable to determine if their ability to create appropriate situation models/inferences differed by proficiency level, or if the intervention compensated for their level of L2 linguistic competence. Standardized test scores on the EPT were used to determine the

level of the ELL students' baseline reading comprehension (tercile split). There were both male and female students representing each proficiency group and experimental versus control groups.

Data Collection

Materials

Students were provided with reading passages followed by comprehension questions that required them to draw pragmatic inferences as well as to answer non-inferential/factual questions. The non-inferential questions were included as baseline data to see if students had a basic level of understanding of the text. The reading passages and comprehension test questions were adopted from already existing materials that prepared ELL students for academic reading (Yaworski, 2006). The experimental group answered an additional adjunct question per passage (See Appendix B).

Procedure

Data were gathered through the following three-step procedure. First, all of the participating students were asked to silently read a short passage for roughly five minutes. Second, half of the students (i.e., the experimental group) explained what the event was about by answering an adjunct question in a written form, while the other half (i.e., the control group) reread the text. The control group was allowed to read each passage a second time as an additional task. Third, all participants were asked to respond to comprehension questions by selecting the correct answer from among three choices per question. This procedure made it possible to analyze student responses according to the two categories of (1) factual/non-inferential information, and (2) pragmatic inferences to see how well each group of students

comprehended discourse requiring different levels of text representation, including, for the pragmatic inference questions, the situation model level.

Data Analysis

To test the previously proposed hypotheses, I investigated how well ELL students performed on reading comprehension questions about the text, which required them to draw inferences, as measured by multiple-choice questions (dependent variable). This was achieved by using mixed analyses of variance (ANOVAs)(i.e., 2 x 2 x 3 design) between subjects for question type (independent variable), whether or not adjunct questioning is implemented as an intervention (the second independent variable), and L2 proficiency ranging from low-intermediate to high intermediate (the third independent variable). By including adjunct questioning as an independent variable, the study was able to investigate the effect of participants' cognitive processing based on their ability to make inferences/develop situation models. The study was designed to explore whether there is a significant interaction between the L2 proficiency factor and the instructional intervention factor. That is, that in order for the intervention to be effective, the students had to have a certain level of L2 proficiency (the linguistic threshold hypothesis). Alternatively, the study might have shown no interaction, yet shown some effect of the intervention factor in which learners of certain proficiency levels benefited from the adjunct questions. A variety of methodologies (including experimental design followed by analysis of students' responses to adjunct questions, including the evidence that they provided from the text) allowed for stronger conclusions regarding the following research questions: (1) To what extent is ELL students' L2 proficiency a factor contributing to their ability to make inferences? (2) Can engaging ELL students in actively explaining target situation models (e.g., by answering explanatory questions about why/how a certain event occurred in the

text as well as what the event caused) compensate for their linguistic competence in the L2? (3)
Does the effect of the intervention depend on L2 proficiency?

Quantitative measures – Reading comprehension task (paper and pencil)

To measure students' level of reading comprehension as a dependent variable, the researcher used reading comprehension tasks. The different levels of proficiency and question types were employed so that the study would reveal a significant main effect of L2 proficiency, indicating that higher proficiency led to greater accuracy across different question types. The researcher predicted that there would be also a significant main effect of question type (i.e., greater accuracy for non-inference questions than for pragmatic inferences) as shown in Rai et al.'s 2011 study. This study resolved whether or not processing information at a pragmatic inference/situation model level was more difficult for less proficient learners.

Qualitative measures - Think-aloud/verbal protocols

As discussed earlier, reading comprehension is a complex process involving both surface level representations (decoding) and underlying text representations (understanding what the writer means). L2 reading requires rather sophisticated skills, including drawing inferences from the text-based representations, which might not be adequately measured by quantitative data such as paper and pencil test scores. Multiple-choice questions and short answer questions have traditionally been used to assess reading comprehension in U.S. school systems. However, Magliano, Millis, Ozuru, and McNamara (2007) argued that students might already know the answers to such comprehension questions without relying on the information they obtain as a result of reading. Therefore, conventional tests did not necessarily measure the depth of comprehension ability, which requires not only the information readers already know but also

text input and cognitive processes (e.g., making connections between old and new information and drawing inferences based on their situation models).

Magliano and colleagues proposed think-aloud protocols as a measure to more accurately assess representations in the reader's mind (Magliano & Millis, 2003; Magliano et al., 2007; Trabasso & Magliano, 1996). Trabasso and Magliano (1996) assessed school-aged children's reading comprehension using the think-aloud method, which focused child readers' attention on each sentence, what it referred to, and how it related to other sentences. Their analysis revealed that the method promoted not only young readers' comprehension but also their monitoring of their own understanding of the text. This in turn supported their global understanding of what they were reading.

Think aloud/verbal protocols are useful for assessing “the nature of inferential processing and strategies that students use in an attempt to understand a given text” (Magliano et al., 2007, p. 116). Thus, the study utilized a variation of the think-aloud/verbal protocol as an dependent variable to assess the readers' ability to explain and infer what the text was about by asking *why* and *how* questions (i.e., adjunct questions) in the context of reading. For example, when asked to explain why/how a certain event occurred in the story at the end, the readers in the experimental group (but not the control group) were prompted to retrieve their memory of the prior text and the thoughts they had associated with the text, which influenced what they comprehended about the text.

By employing this combined assessment tool in my L2 reading comprehension study, the researcher was able to assess how well ELL students are able to maintain coherence by making connections with prior discourse through situation models. Just as more skilled and less skilled readers performed in Trabasso and Magliano's (1996) study, ELL readers with an high-

intermediate level of English proficiency should have been able to maintain coherence throughout the text, while intermediate and/or low-intermediate level ELL students were likely to discuss about sentences in isolation without making any connections to the prior discourse. The findings from this study provided evidence that the depth of underlying knowledge representations used by high-intermediate, intermediate and low-intermediate level L2 readers varied and affected their reading comprehension. Additionally, the findings suggested that not all readers were able to make connections across different parts of the discourse when adjunct questions were used as a strategy to raise their metacognitive awareness.

Reliability and validity

According to Goodwin (2010), reliability refers to “the extent to which measures of the same phenomenon are consistent and repeatable” (p. 569). In order to minimize measurement errors, researchers must use measurements that produce more or less the same outcome such as standardized test scores. Without relying on a consistent measure, we cannot determine what a particular score actually means. As mentioned in Chapter One, EPT scores were used to determine the L2 proficiency level of participating readers. Those who scored 50 on the EPT were very unlikely to score 55 if tested for a second time, say, a week later. Therefore, no matter how many times being tested, the students’ EPT scores should not have varied if taken within a similar time frame (e.g., when testing a week later). This allowed me to compare across proficiency levels that were reliably different to examine the participants’ ability to draw inferences at a situation model level.

The reliability of this study was also increased by using the stimuli (i.e., reading passages followed by comprehension questions) that were consistent in length and in the number of inferential, and non-inference questions included; testing each type of inference multiple times

also ensures greater reliability. Participants read a total of five one-paragraph long passages. Each paragraph consisted of an average of four sentences. To maintain coherence, the lengths of individual paragraphs varied from four to six sentences. Each passage was followed by one non-inferential/factual question and two pragmatic inference questions. This uniformity across the stimuli enabled me to equate the measurements across levels of the different independent variables (i.e., three levels of L2 proficiency and two different treatments), which could be tested via a pilot study.

Use of the think aloud/verbal protocol as a qualitative data collection tool allowed me to triangulate the data, which further increased the reliability of this study. When evaluating the quality of participants' responses to see if they were targeting appropriate situation models, I asked another trained think-aloud protocol grader to score the responses to establish inter-rater reliability (i.e., as measured by Cohen's Kappa). In addition, the reliability of the various measures used in this study was examined via split-half reliability or Cohen's Kappa.

Having established the reliability of each of these measures, its validity has to be taken into consideration. Validity refers to the extent to which a measure of some construct truly measures "what it has been designed to measure" but not other constructs (Goodwin, 2010, p. 131). For example, a valid measure of the effect of a reading comprehension intervention should not measure something other than that, such as driving ability. Validity assumes reliability. As such, "valid measures must be reliable" though the opposite is not true (Goodwin, 2010, p. 134).

By using a matched control group, which was identical to the experimental group except that it was not exposed to the experimental variable (i.e., responding to adjunct questions), I equated the two groups on various extraneous variables that might otherwise affect the results (e.g., sex, and baseline proficiency level). Matching is particularly important when random

assignment is not possible so that you can control extraneous variables that you are not interested in (e.g., gender and baseline language proficiency level). Once matched pairs were formed (i.e., same sex, and the closest matching EPT scores), participants from each pair were randomly assigned to the experimental and control conditions. The same applies to EPT scores; both control and experimental groups consisted equally of low-intermediate, intermediate, and high-intermediate leveled ELL students as measured by their performance on the EPT. This matching increased the internal validity of the proposed study (i.e., the degree to which the manipulated variable makes a difference).

Within internal validity, perhaps the most crucial role is played by construct validity, namely the degree to which a test adequately measures the construct (Goodwin, 2010). One way of assessing construct validity in this proposed study was to see if the comprehension questions were related to one another when they were supposed to be related (referred to as convergent validity) and were unrelated when they are supposed to be unrelated (referred to as discriminant validity). For example, both types of reading comprehension questions, pragmatic inference and factual (non-inference) questions, should have been more correlated to other questions of the same type (inferential) than they were to questions of the other type (non-inferential), across the different reading passages. (By analogy, to the extent that reading and math questions in the SAT have construct validity, two reading comprehension questions within the SAT should be more highly correlated to each other than a reading question and a math question in the SAT and vice versa.) If this were true, then the reading comprehension measure used in the study can be assumed to have a certain degree of construct validity.

In contrast to internal validity, external validity refers to the extent to which research findings generalize outside the particular context of the study you are conducting (Goodwin,

2010). The external validity of a study is determined by the degree to which the results of the study are applicable to other populations, environments, and times beyond those of the experimental setting. If the proposed study has a high degree of external validity, the findings should apply to other people, places, and times. In theory, such generalizability is realized only when the study is replicated in various research settings. However, Pressley and Harris (1994) argue that external validity can be increased by modifying the context of educational intervention research; such modification, for example, might involve “doing reading interventions in actual reading groups rather than with groups of children brought together for the first time by an experimenter” (p. 196). Data collection of the study took place in non-artificial settings (e.g., students’ English language/professional development classes). Thus, the results from the questioning/explaining intervention as reported in Chapter Four were likely to be generalizable beyond the context of the study.

Chapter 4 - Data Results and Analysis

This study was designed to answer the following overarching research questions:

- (1) To what extent is ELL students' ability to make L2 inferences affected by their L2 proficiency? (Interaction between L2 proficiency and question type)
- (2) Can engaging ELL students in actively explaining target situation models (e.g., why and how a certain event occurred in the text as well as what the event caused) improve their L2 comprehension? (Main effect of intervention on comprehension)
- (3) Does the effect of the above intervention on L2 comprehension depend on learners' L2 proficiency? (Interaction between intervention and L2 proficiency on comprehension)

Based on the above three research questions, the following four hypotheses regarding main effects and interactions were tested:

- (1) Factual/non-inferential questions are more likely to be answered appropriately by L2 learners than those that require pragmatic inferences based on the reader's situation models (The main effect of the question type on comprehension)
- (2) Pragmatic inferences are more likely to be made by the learners with higher proficiency than those with lower proficiency, and the factual/non-inferential questions will be correctly answered at equal rates by learners of all levels of L2 proficiency. (The interaction between the question type and L2 proficiency on comprehension)
- (3) As ELL students actively engage in responding to explanatory/adjunct questions, they will better create situation models, and thus better comprehend the text. (The main effect of the intervention on comprehension)

- (4) The effectiveness of having ELL students respond to explanatory/adjunct questions will differ according to proficiency level. (The interaction between the intervention condition and L2 proficiency on comprehension)

The situation model theory (Kintsch, 1998/2007), the event indexing model (Zwaan & Radvansky, 1998), and the linguistic threshold hypothesis (Anderson, 1984; Carrell, 1991; Clarke, 1980) served as the theoretical frameworks for the educational intervention used in this study. As discussed in Chapter One, the three theoretical frameworks include the ideas that: (a) skilled readers are able to construct a situation model to maintain a coherent understanding of the text that they are reading, (b) readers use event indices as part of their text representation to monitor their understanding of the causes of events and the protagonist's intentions, and (c) ELL readers' L2 proficiency affects their ability to use higher-order reading strategies to improve their L2 reading comprehension, including using situation models to infer causal connections among events described in a passage.

An experiment involving a reading instruction intervention was carried out; both qualitative and quantitative data were collected. The researcher investigated the impact on L2 comprehension of having learners actively create situation models during L2 reading. The possible role of readers' L2 English proficiency levels on the impact of the instructional intervention also was explored. The data were analyzed according to the methodology outlined in Chapter Three. This chapter reports the results found and the analyses carried out for this study. The chapter includes: (a) a description of the sample population, (b) a description of the measures, (c) a description of the quantitative analysis, including the descriptive statistics, the inferential statistics (e.g., ANOVAs), and the results of the tested hypotheses, and (d) a qualitative representation of readers' thoughts generated during the reading task.

Description of the Sample Population

The sample population who participated in this study consisted of 117 English language learners. The vast majority of these participants teach English as a foreign language in their home country, Ecuador. Thus, presumably none had low proficiency in English. There were 40 male students and 77 female students, reflecting the general gender bias in the teacher population of Ecuador. They varied in age from 23 to 50 with an average age of 32, though as the following table shows, the modal age was slightly younger (28) than the mean (32) and median (31). All of the participants were Spanish L1 speakers who were studying English as their L2 for academic purposes. The sample population was thus homogenous in terms of their profession, nationality, and native language.

Table 4.1 Age of Sample Population

Age Description	Value
Mean Age	32.28
Modal Age	28
Median Age	31
Standard Deviation	5.33
Minimum	23
Maximum	50

The participants were enrolled in either a seven- or nine-month English language/professional development program (starting between Fall 2012 and Spring 2013, depending on the time of their arrival at their partner university in northeastern Kansas). The

participating students' overall English language proficiency levels were determined by a standardized measurement (the English Proficiency Test, or EPT), which was created and administered by the partner university. Through random assignment, half of the sample population was placed in the experimental group (58), while the other half was placed in the control group (59). The 117 participants were divided into three approximately equal-sized proficiency groups: low-intermediate (38), intermediate (39), and high-intermediate (40). However, the random assignment of participants to intervention conditions resulted in there being more intermediate students in the experimental group (24) than in the control group (15). The distribution of the 117 participants in each condition as a function of L2 proficiency level is summarized in Table 4.2.

Table 4.2 Proficiency of Sample Population

Condition	L2 Proficiency (EPT)	Population
Control (59)	Low-intermediate (Bottom 3 rd)	23
	Intermediate (Middle 3 rd)	15
	High-intermediate (Top 3 rd)	21
	All levels	59
Experimental (58)	Low-intermediate (Bottom 3 rd)	15
	Intermediate (Middle 3 rd)	24
	High-intermediate (Top 3 rd)	19
	All levels	58
Total		117

The study was designed to investigate the effects on L2 learners' reading comprehension of an instructional intervention that either did or did not encourage them to create situation models while reading (a between-groups factor), and a possible moderator of the effectiveness of the intervention, namely the learners' L2 proficiency levels (a subject factor). Specifically, participants in the experimental condition received a treatment that involved responding to adjunct questions while reading. This process elicited their explanations of underlined situations, which required them to infer causal relationships in the text. Participants in the control condition simply reread the text rather than explaining it. All participants were then asked to answer multiple-choice comprehension questions, which included both inferential and non-inferential questions (a within-group factor). Participants were then compared in terms of their performance on the reading comprehension task in the two intervention conditions for each of the different proficiency levels (low-intermediate, intermediate, and high-intermediate).

Description of the Measures

The quantitative data collection measures used in this study included: (1) a reading task including five reading passages with three comprehension questions per passage (i.e., 15 questions total) and (2) the students' EPT scores. The reading task consisted of five different ecologically valid college-level text passages that were derived from an English for academic purposes reading textbook. Each passage was followed by three multiple-choice comprehension questions. The three questions consisted of two inferential questions and one non-inferential (factual) question for each passage, in order to hold constant the within-group factor of question type. Furthermore, the passage order (1-5) was randomized so that performance would not be affected by the order in which participants read and responded to the passages. This random arrangement also prevented neighboring participants from copying each other's answers.

Prior to answering the set of comprehension questions for a given reading passage, the participants in the experimental condition had an additional task. They were required to explain an underlying situation in the passage by responding to an *adjunct question* (e.g., “What was Wilkie’s intention of telling his audience that if they reelected Roosevelt, they would expect war in 1941?”) (Anderson et al., 1977; Kintsch, 2005). This task encouraged participants in the experimental condition to make an inference about the protagonist’s intentions/goals, which were not explicitly stated in the text. Other adjunct questions were designed to elicit inferences regarding causal relationships embedded in passages (e.g., “What was the outcome of Kennedy’s mentioning the risk of nuclear war?”). Such adjunct questions were used as an experimental intervention based on the idea that they would engage the readers in trying to determine causality and intentionality, resulting in a better situation model for the text, and thus better comprehension of it.

The students’ previously recorded EPT scores were used to measure their English L2 proficiency, including their proficiency in listening, reading, and structure/grammar. Each EPT section contained fifty multiple-choice questions, which the students were given twenty-five minutes to answer. The EPT scores were reported as T-scores (i.e., a student’s Z-score * 10 + 50), which meant that each student’s performance was on a normative scale relatively equivalent to other established normative English L2 proficiency measures (e.g., TOEFL). The EPT test is normally administrated at the beginning, middle, and end of the program. By the time this data collection took place in February and March 2013, the participants who had been in the nine-month program had already taken the first two tests, whereas those who were in the seven-month program had only taken the first test. Therefore, the latest EPT scores were selected for this study (i.e., the middle scores from the nine-month program and the beginning scores from the

seven-month program). The participants' EPT scores were the sum of their scores for the listening, reading, and structure/grammar sub-tests, and these sums varied from 101 to 191. After this, the distribution of summed EPT scores was divided into three proficiency levels by tertile split: the bottom third (101-147) called the low-intermediate level, the middle third (148-160) called the intermediate level, and the top third (161-191) called the high-intermediate level.

The measure used for the qualitative analysis was a variation of a think-aloud protocol. As described in Chapter Three, the thoughts generated by the experimental group were collected in a written form in order to assess their ability to infer and explain the underlying causal or intentional structure of the text. The participants' responses to each of the five adjunct questions were transcribed and coded into one of the following categories: (1) overarching ideas (causality and intentionality), (2) pragmatic knowledge/knowledge of the world to infer meaning, and (3) coherence. These categories were based on the taxonomy of constructionists' classification of inferences, including superordinate goals, causal antecedents, and global thematic inferences (Graesser et al., 1994), adapted for the current study.

Under the relevant categories, the participants' responses were further coded into one of the three possible subcategories depending on the extent to which they understood the overarching ideas, used their pragmatic knowledge, or showed coherent interpretations of the text. Thus, assigning each participant response to one of the nine subcategories allowed the researcher to put the qualitative data into a systematic shape involving quantification (Wragg, 2012). Wragg argues that quantifying categories provides classroom research a degree of objectivity. Table 4.3 summarizes how a total of nine subcategories were organized within the three major categories. Once the think-aloud data were coded by the author, a second coder was

asked to code the same responses in order to establish inter-rater reliability (i.e., Cohen’s Kappa). In the examples below, coding for each subcategory is explained (see the italicized comments).

Table 4.3 Coding for Qualitative Data

Category	Subcategories	Examples
1. Overarching ideas (Example Question from passage # 2: “What was Wilkie’s intention of telling his audience that if they reelected Roosevelt, they would expect war in 1941?”)	a. Understanding causal relationships and/or a protagonist’s intention/goal	“Wilkie’s intention was to convince people not to vote for Roosevelt by making them feel afraid of being part of a war.” (Participant #7) – <i>The participant understood the protagonist Wilkie’s main goal of his speech.</i>
	b. Superficial understanding of causal relationships or a protagonist’s intention/goal	“Wendell Wilkie charged that Franklin wanted to take the US into war but Roosevelt didn’t explain if [sic] will be war, but he promised not to send young people to fight into foreign wars.” (Participant #20) – <i>While the participant understood what Wilkie had done, s/he did not explain why he had done it.</i>
	c. Lack of understanding of causal relationships or a protagonist’s intention/goal	“Wilkie pretended to make citizens change the [sic] opinion in order to [sic] Franklin D. Roosevelt lost [sic] the election and he didn’t become a President.”

		(Participant #121) – <i>The participant seemed to have misused the term “pretended,” which canceled the meaning of Wilkie’s attempt to convince the audience not to vote for Roosevelt.</i>
<p>2. Pragmatic knowledge/knowledge of the world</p> <p>(Example Question from passage # 5: “What was the outcome of Kennedy’s mentioning the risk of nuclear war?”)</p>	a. Use of appropriate pragmatic knowledge	“Kennedy’s outcome [sic] is to inform the American people about the missiles. The impact of them, and provide quarantines to keep their nation safe.” (Participant #3) – <i>The participant understood Kennedy had prepared Americans to mobilize for the Cuban missile crisis.</i>
	b. Marginally appropriate use of pragmatic knowledge	“He wanted the missiles were [sic] removed.” (Participant #64) – <i>The participant understood what Kennedy wanted but misunderstood the term “outcome.”</i>
	c. Lack of pragmatic knowledge	“Joh F. Kenedy want [sic] to avoid the cold war.” (Participant #85) – <i>The participant was unaware that the US and the Soviet Union were already in the Cold War, and Kennedy gave his speech</i>

		<i>during the middle of it.</i>
<p>3. Coherence</p> <p>(Example Question from passage # 1: “Why did the dogs follow the ladies?”)</p>	<p>a. Made connections to create coherent interpretation of the text</p>	<p>“The ladies were carrying food inside the bags, dogs smelled the food, and they wanted to try this. That is why many dogs were following the ladies.” (Participant #2) – <i>The participant provided a coherent summary of the story throughout the response.</i></p>
	<p>b. Moderately coherent interpretation of the text</p>	<p>“Because the dogs wanted to have a great banquet [sic] of food for that holiday, they smelled the hams and they just wanted to eat that food.” (Participant #16) – <i>Only the second half of the participant’s response was coherent. The first half showed a misunderstanding of who would enjoy the holiday meal (unless the participant was trying to be humorous).</i></p>
	<p>c. Lack of basic coherence: incoherent comprehension of the text, irrelevant details, lack of lexical/text knowledge to comprehend the text</p>	<p>“There weren’t any dogs, they are mentioned because they were in a rush.” (Participant # 73) – <i>The participant’s response was incoherent and provided details irrelevant to the story.</i></p>

Analysis of Quantitative Data

The mean reading comprehension task scores for each condition (control or experimental) across proficiency levels (low-intermediate, intermediate, and high-intermediate) were statistically analyzed. This was to investigate: 1) the effect of the experimental instructional intervention (i.e., explaining the underlying causal or intentional structure of the text by answering adjunct questions) on L2 readers' comprehension and 2) whether the effectiveness of the intervention depended on the learners' proficiency level.

First, each of the following factors was analyzed using a 2 (Question Type: Factual vs. Inferential [within-subjects]) x 2 (Intervention Condition: Experimental vs. Control [between-subjects]) x 3 (L2 Proficiency: Low-intermediate vs. Intermediate vs. High-intermediate [between-subjects]) Mixed ANOVA. Table 4.4 shows in the following order: (1) the main effect of the question type (factual vs. inferential), (2) the interaction between the question type and intervention condition, (3) the interaction between the question type and L2 proficiency, and (4) the interaction between the question type, intervention condition, and L2 proficiency.

Table 4.4 Summary of the 2 x 2 x 3 Mixed ANOVA (Test if Within-Subjects Effects and Within-Subjects x Between-Subjects Effects) Coding for Qualitative Data

Effect	Type III SS	Df	Mean Square	F	Sig.
Q type	0.26	1	0.26	8.649	0.004
Q type x Condition	0	1	0	0.009	0.926

Q type X	0.069	2	0.034	1.143	0.323
Proficiency					
Q type X	0.059	2	0.03	0.989	0.375
Condition X					
Proficiency					
Error	3.336	111	0.03		

*Statistical significance alpha level = .05

As shown in Table 4.4, there was a strong and statistically significant effect of question type ($F(1, 111) = 8.649, p = .004$), such that, overall, participants did better with the non-inferential (factual) questions than they did with the inferential questions. This finding both confirms the first hypothesis, namely that there is a main effect of question type, and is consistent with the results of earlier studies (Hammadou, 1991; Horiba, 1996; Kembo, 2001; Rai et al., 2011; Shimizu, 2005).

Table 4.4 also answers the first research question: To what extent is ELL students' L2 proficiency a factor contributing to their ability to make appropriate inferences? This question formed the basis of the second hypothesis, namely that there would be an interaction between the question type and proficiency. As illustrated in Table 4.4, the finding of a non-significant trend disconfirms the second hypothesis that there would be interaction between the question type and proficiency. These results indicated a non-significant interaction of question type with L2 proficiency ($F(2, 111) = 1.143, p = .323, n.s.$). Yet, there is a tantalizing trend of the interaction. Figure 4.1 shows a hint of an interaction between these two variables, in that the low-intermediate and intermediate learners did much better on the factual than the inferential questions (Low-int: $t [37] = 1.943, p = .060$; Int: $t [38] = 3.477, p < .001$), whereas the high-

intermediate learners did well on both factual and inferential questions ($t [39] = 0.473, p = .639, n.s.$). The pattern shown by three t-tests suggests that the non-significant interaction between question type and L2 proficiency likely would have been significant if the difference between factual and inferential questions for the low-intermediate learners ($p = .060$) had been slightly larger. More specifically, Figure 4.1 shows that from the low-intermediate to the intermediate level, the L2 learners made more improvement on the factual questions than on the inferential questions. However, from the intermediate to the high-intermediate level, the L2 learners improved on both question types, with more marked improvement on the inferential questions. Thus, learning across proficiency levels seems to show different patterns for the different question types.

Figure 4.1 Comprehension Accuracy across Question Type and L2 Proficiency

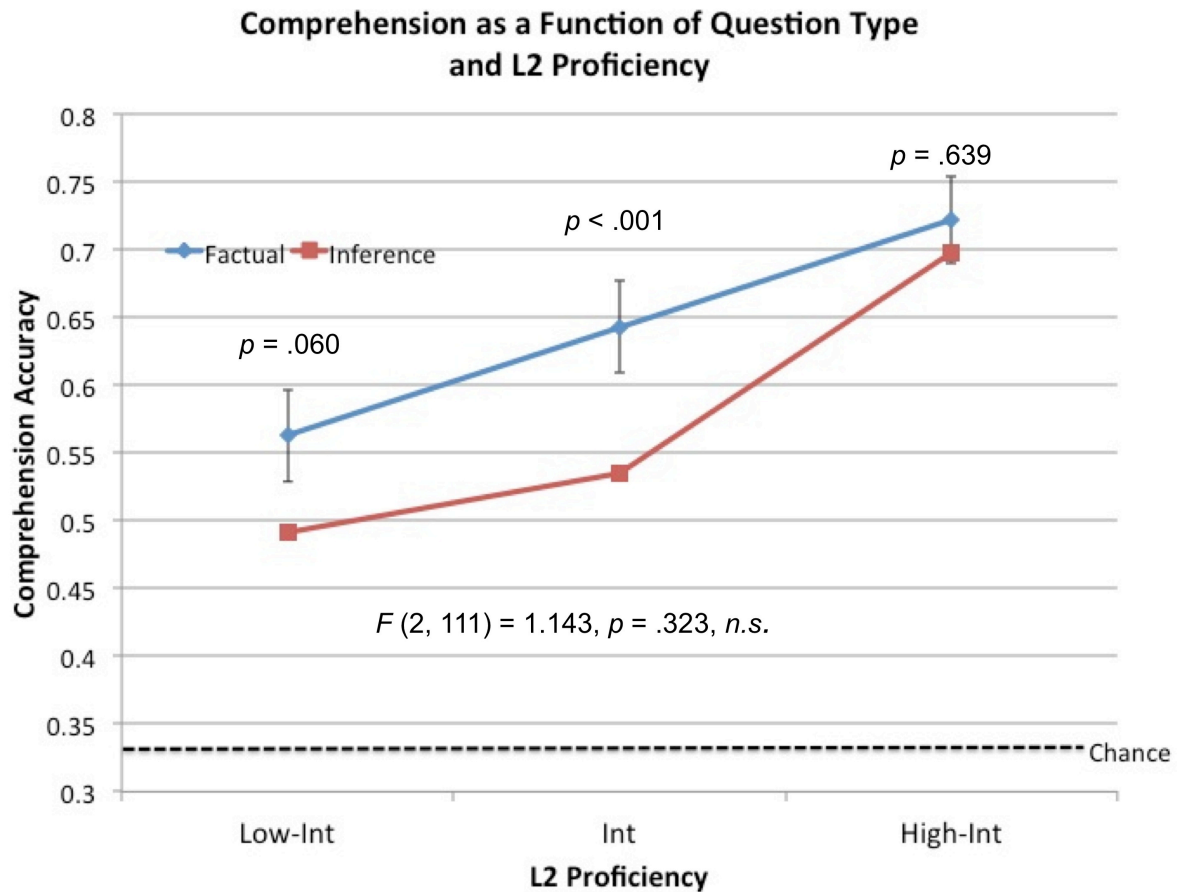


Table 4.4 also depicts a statistically non-significant two-way interaction between question type and the intervention condition ($F(1, 111) = 0.009, p = .926, n.s.$), such that learners of both conditions show basically the same effect of question type. Similarly, there was a non-significant three-way interaction among question type, intervention condition, and L2 proficiency ($F(2, 111) = 0.989, p = .375, n.s.$), such that the relationship between question type and proficiency was unchanged by the intervention condition.

The between subjects factors were intervention (experimental vs. control conditions) and L2 proficiency (low-intermediate, intermediate, and high-intermediate). As shown in Table 4.5, overall, participants in the intervention group performed moderately better than the non-intervention group. The same was true with L2 proficiency. Participants in the high-intermediate

level in either condition performed better than their intermediate and low-intermediate peers at a moderate scale that is continuous. Table 4.6 shows the main effect of each of these factors and the interaction between them.

Table 4.5 Summary of Descriptive Statistics for Comprehension Accuracy as a Function of Intervention Condition and L2 Proficiency

				95% Confidence Interval	
Condition	Proficiency	Mean Accuracy %	Std. Error	Lower Bound	Upper Bound
Control	Lo-Int	0.517	0.027	0.463	0.572
	Int	0.587	0.034	0.52	0.654
	Hi-Int	0.686	0.029	0.629	0.742
Experimental	Lo-Int	0.537	0.034	0.47	0.604
	Int	0.592	0.027	0.539	0.645
	Hi-Int	0.735	0.03	0.675	0.794

Table 4.6 Summary of the 2 x 3 ANOVA (Tests of Between-Subjects Effects)

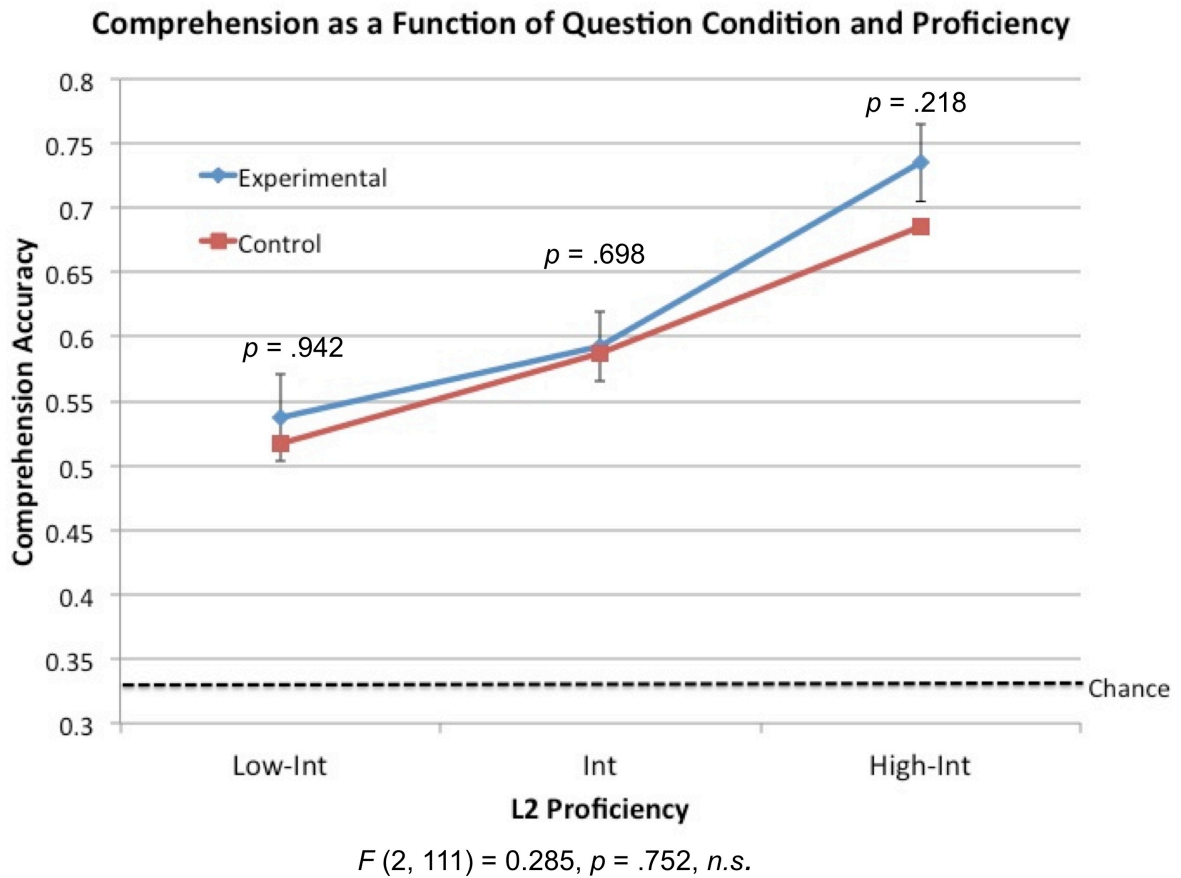
Effect	Type III SS	Df	Mean Square	F	Sig.
Condition	0.034	1	0.034	0.984	0.323
Proficiency	1.334	2	0.667	19.392	0.000
Condition X Proficiency	0.02	2	0.01	0.285	0.752
Error	3.818	111	0.034		

*Statistical significance alpha level = .05

The only statistically significant effect was that of L2 proficiency on overall comprehension accuracy ($F(1, 111) = 19.392, p = 0$), such that higher proficiency L2 learners were more accurate on all comprehension questions in general, as expected. More importantly, Table 4.6 answers the second research question: Can engaging ELL students in actively explaining target situation models compensate for lack of linguistic competence in the L2? Specifically, there was a statistically non-significant main effect of the intervention ($F(1, 111) = 0.984, p = .323, n.s.$), which is inconsistent with the third hypothesis, that ELL students' active engagement in responding to explanatory questions would lead to creating situation models, thus better comprehension.

Table 4.6 also answers the third research question: Does the effect of the intervention depend on learners' L2 proficiency? Specifically, there was a non-significant interaction between the intervention factor and L2 proficiency ($F(2, 111) = 0.285, p = .752, n.s.$), which is inconsistent with the fourth hypothesis, that an activity designed to encourage L2 readers to ask adjunct questions while reading would lead to better comprehension only for those with enough L2 proficiency. However, Figure 4.2 shows a slight tendency towards such an interaction. Specifically, as Figure 4.2 illustrates, the greatest mean difference between the two intervention conditions seems to be among the high-intermediate students (0.686 for the control group vs. 0.735 for the experimental group, High-Int: $t [38] = 1.251, p = .218, n.s.$), whereas lower two proficiency groups showed virtually no difference at all (Low-int: $t [36] = .073, p = .942, n.s.$; Int: $t [37] = .392, p = .698, n.s.$). This suggests that the adjunct questions may have been slightly more helpful for the highest proficiency group, which, if true, would be consistent with the linguistic threshold hypothesis and earlier findings by Carrell (1991) and Nassaji (2007). The possible implications of these quantitative data will be further discussed in Chapter Five.

Figure 4.2 Comprehension Accuracy Summary across Conditions and L2 Proficiency



Qualitative Analysis of Readers’ Thoughts (Think Aloud)

The frequency of the experimental condition participants’ responses to each of the five adjunct questions was determined after the raters assigned each response to one of the above nine subcategories (refer to Table 4.3). Then the information was organized by proficiency level as a within-group factor. Pearson’s correlation coefficient was used to see if the nature of the participants’ responses would be a good predictor of their performance on the think aloud questions. For example, those who were in the high-intermediate proficiency level should have been able to provide responses that were coded as “made connections to create coherent

interpretation of the text” significantly more often than intermediate and low-intermediate proficiency participants. Likewise, the participants with low-intermediate proficiency should have provided responses that were likely to be coded as “lack of basic coherence” more frequently than intermediate and high intermediate proficiency participants.

The fifty-eight participants in the experimental condition were asked to generate their thoughts about all five passages by responding to the adjunct questions below. They provided an explanation for each question while reading. This took place prior to their responding to a set of three comprehension questions per passage, including one factual and two inferential questions. The rationale for this task was that answering adjunct questions about causal relationships and intentionality (why and how questions) would lead to participants creating situation models during the process of reading. In theory, treatments such as this should help participants make inferences about the underlying meaning (i.e., causal or intentional structure) of the text, as inferences require situation models for a coherent understanding (Graesser, 2007; Graesser, Singer, & Trabasso, 1994). The adjunct questions (one for each passage) included:

- Why were dogs following the ladies? (Causal relationship)
- What was Wilkie’s intention in telling his audience that if they reelected Roosevelt, they would expect war in 1941? (Intentionality)
- What are the effects of consuming high doses of vitamins? (Causal relationship)
- What is the purpose of landfill linings? (Intentionality)
- What is the outcome of Kennedy’s mentioning the risk of nuclear war? (Causal relationship)

(Note: The full text passages and their corresponding comprehension questions, aligned with the above adjunct questions, are located in Appendix B.)

As described previously, the intervention procedures were based on a variation of the think aloud protocol with which participants' ability to make inferences about the central idea of the text was assessed. After coding the participants' responses to the adjunct questions into the subcategories shown in Table 4.3, the researcher took a frequency count according to each subcategory across the three levels of L2 proficiency as a within-group factor.

Of the fifty-eight participants in the experimental condition, fifty-seven students responded to all five questions, which resulted in a total of 285 different responses. A total of two graders (including the researcher) who hold a master's degree in English and applied linguistics were involved in this qualitative analysis. After the other grader was trained in assigning participant responses to the categories, both graders coded individually all 285 responses. The raters were blinded to the proficiency condition in order to avoid possible biases for/against high/low proficient participants. After the coding was completed, the graders discussed the process to clarify any ambiguity that they had faced. For example, the following response could well be categorized into 1-a (understanding causal relationships/a protagonist's intention) and/or 2-a (use of appropriate pragmatic knowledge).

Wilkie's intention was [sic] make people think against Roosevelt in the same way people is [sic] against wars. Because Wilkie wanted to be [sic] president. (Participant # 2)

In fact, the co-grader originally coded this response mostly for 1-a but partially for 2-a. In order to increase their degree of agreement, the graders came to an agreement in defining the classification (operational definition) so that they could uniformly recode the response as one score, which was 1-a. Although the response could be causal as well as pragmatic, the graders

regarded it as causal due to the use of the logical connector “because.” Other ambiguous items were treated in a similar fashion. Thus, every response was given only one score by each rater. Of the 285 items, the two raters agreed on initial classification of 198 items, which was equal to 69% inter-rater agreement.

Table 4.7 summarizes the raw frequencies of participants’ categorized think aloud responses. For example, consider the frequency of think aloud response ratings for the High-intermediate participants shown in Table 4.7. These reflect ratings of the responses for 19 such participants in the experimental condition, for each of the 5 passages’ adjunct questions, for both of the 2 raters, for a total of $19 \times 5 \times 2 = 190$ ratings. By contrast, Table 4.8 provides the frequency as a percentage of the total responses for a given language proficiency level, because the number of participants in each level varies, ranging from 14 to 24. For example, the percentages of think aloud responses for the High-intermediate participants in Table 4.8, reflects the totals for each of the cells in the first row of Table 4.7 divided by the row total of 190 ratings.

Table 4.7 Raw Frequencies of Think Aloud Responses

Subcategory	Overarching Ideas			Pragmatics			Coherence			Total
	1a	1b	1c	2a	2b	2c	3a	3b	3c	
High-Intermediate	85	15	5	1	15	18	19	23	9	190
Intermediate	87	17	14	12	20	22	11	43	14	240
Low-Intermediate	36	13	7	3	5	15	11	32	18	140
Total	208	45	26	16	40	55	41	98	41	570

Table 4.8 Percentage of Frequencies of Think Aloud Responses

Subcategory	Overarching Ideas			Pragmatics			Coherence			Total %
	1a	1b	1c	2a	2b	2c	3a	3b	3c	
High-Intermediate	44.7	7.9	2.6	0.5	7.9	9.5	10	12.1	4.7	100
Intermediate	36.3	7.1	5.8	5	8.3	9.2	4.6	17.9	5.8	100
Low-Intermediate	25.7	9.3	5	2.1	3.6	10.7	7.9	22.9	12.9	100

We will consider the categorical data for the high-intermediate level participants first. As shown in Table 4.8, the row for High-intermediate participants shows that among the 9 subcategories of responses, the single most frequent subcategory (44.7%) was 1-a, within the “Overarching Ideas” category. Sub-category 1a was “understanding causal relationships and a protagonist’s intentions/goal.” The second most frequent subcategory (12.1%) was 3-b, having moderately coherent interpretation of the text, which was followed by and 3-a (10%), making connections to create coherent interpretation of the text.

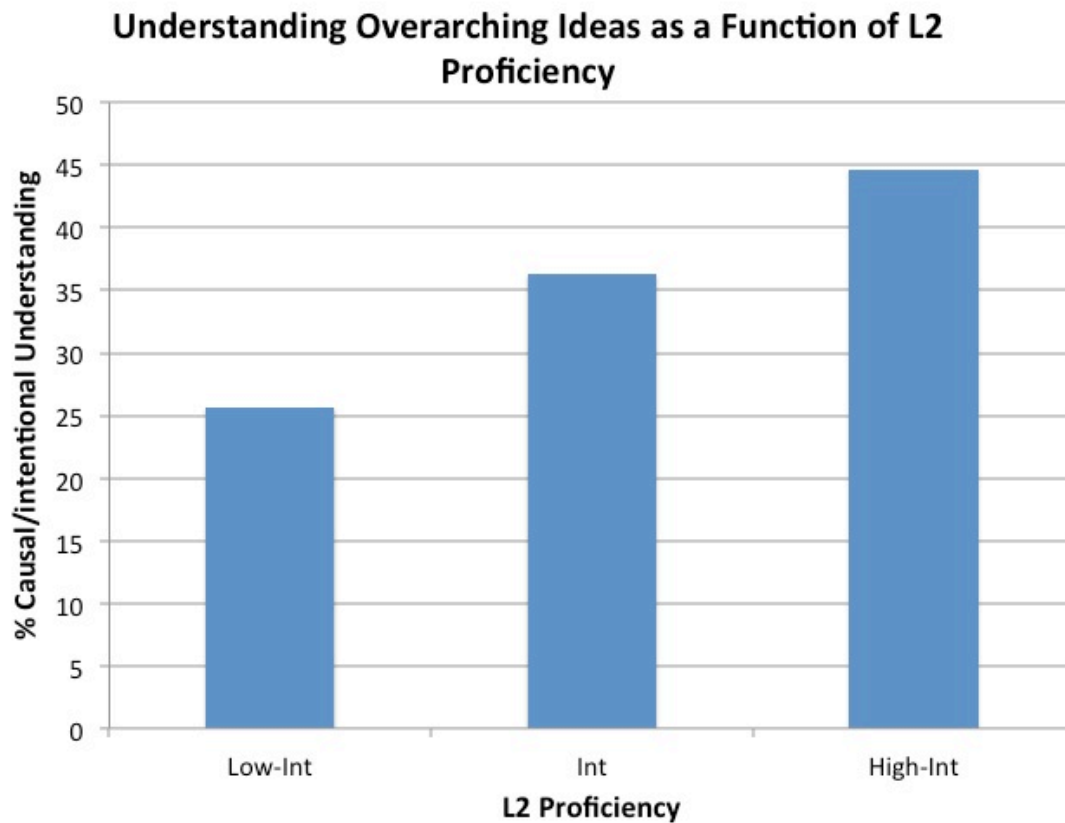
Next, consider the rated responses of the intermediate level participants. Similarly to the high-intermediate participants, the most frequent sub-category (36.3%) was 1-a. Also like the high-intermediate group, the second most common subcategory for this group (17.9%) was 3-b . However, the third most common subcategory (9.2%) was 2-c, lack of pragmatic knowledge.

Among the low-intermediate level participants, 1-a was again the most common subcategory (25.7%). Yet, unlike the results for the higher proficiency level participants, this subcategory did not represent the vast majority of the total responses. In fact, the frequency of the second most commonly coded response (22.9%), for 3-b, was very close to that of 1-a.

Regarding low-intermediate level responses, it is noteworthy that 3-c, lack of basic coherence (12.9%), was the third most frequent subcategory, whereas among the high-intermediate participants it was the least frequent.

The overall tendency indicated that high-intermediate and intermediate-level participants understood causal relationships and protagonists' intention/goal (1-a) embedded in the text more than their low-intermediate-level counterparts. Of the 208 coded responses across proficiency levels in this subcategory, 85 (44.7%) and 87 (36.3%) responses were generated by the high-intermediate and intermediate groups respectively, whereas only 36 (23.7%) responses were reported by the low-intermediate-level participants. Figure 4.3 visually represents this linear relationship between participants' understanding of the causal/intentional dimension, as determined by the raters, and the participants' L2 proficiency level. Thus, the causality and intentionality aspect of this think aloud data provided suggestive evidence for the linguistic threshold hypothesis (Anderson, 1984; Carrell, 1991; Clarke, 1980), which is related to the third research question (i.e., Does the effect of the intervention depend on L2 proficiency?).

Figure 4.3 Level of Understanding Causal/Intentional Dimension across L2 Proficiency



No such linear relationship was observed for other major categories. Nevertheless, this data supported the idea that the participants' level of L2 proficiency (i.e., the ELP measured proficiency) predicted the quality of their responses in terms of causality and intentionality., especially the high-intermediate group, although Figure 4.3 shows a continuous decline across the three groups. In other words, the higher the L2 proficiency, the more they generated causal and intentional inferences accurately.

Table 4.9 presents the think aloud responses coded by both raters that fell in the top four subcategories for each of the three L2 proficiency levels. The selected subcategories are listed in order by the frequency of total responses assigned. A response example from each subcategory is

provided to represent what a thought process is like among participants in the experimental condition across different proficiency levels.

Table 4.9 Top 4 Think Aloud Responses across L2 Proficiency

Proficiency	Subcategories	Think Aloud Response Examples
1. High-intermediate	1-a. Understanding causal relationships and/or a protagonist's intention/goal	“There are a [sic] bad effects in consuming high doses of vitamins. One of this [sic] is a liver damage, another one could be headaches, etc. For a healthy life, people must have a good and normal diet.” (Participant #45)
	3-b. Moderately coherent interpretation of the text	“Modern waste disposal land fill [sic] create oxygen free conditions that reduce bacterial and fungal action, which decompose material rapidly.” (Participant #21)
	3-a. Made connections to create coherent interpretation of the text	“Landfill linings are used. So, they create oxygen-free conditions to slow down bacteria and fungi. <u>For these reasons</u> , newspapers and hotdogs were not destroyed.” (Participant #58) <i>Note: The participant used appropriate transition words.</i>
	2-c Lack of pragmatic knowledge	“Landfill linings’ purpose is to reduce the great amount of bacterias [sic] and fungi by giving enough oxygene [sic].” (Participant #99)
2. Intermediate	1-a. Understanding causal relationships and/or a protagonist's intention/goal	“I think that the effects of consuming high doses of vitamins can produce a lots [sic] of diseases and terrible damage in your liver, for example, like vitamin D is the most potentially toxic vitamin.” (Participant #49)
	3-b. Moderately coherent interpretation of the text	“The purpose is to prevent damage or waste on land avoiding oxygen-free conditions to decrease fungus doing their actions to destroy things to convert in other material.” (Participant #67)
	2-c. Lack of pragmatic knowledge	“Wilkie's intention was to see if Roosevelt were

		reelected, he would keep his promise of not sending boys into any foreign wars even though the <u>United States was taken into the war.</u> ” (Participant # 33)
	2-b. Marginally appropriate use of pragmatic knowledge	“Wilkie's intention was to avoid the United States break [sic] into a new world war and prevent that American solders and people get involved on [sic] it with fatal consequences.” (Participant #5)
3. Low-intermediate	1-a. Understanding causal relationships and/or a protagonist’s intention/goal	“To consume high doses of vitamins can cause damage in the body organs.” (Participant #120)
	3-b. Moderately coherent interpretation of the text	“The purpose of landfill linings is to reserve or keep something that can be recognize [sic] after a long time.” (Participant #108) <i>Note: The participant understands only the secondary purpose of landfill linings.</i>
	3-c. Lack of basic coherence: incoherent comprehension of the text, irrelevant details, lack of lexical/text knowledge to comprehend the text	“If the same president is reelected, he may expect war in April. The presidents [sic] want [sic] said [sic] before that they [sic] boys are not going to be sent into any wars.” (Participant #78)
	2-c Lack of pragmatic knowledge	“The outcome of Kennedy’s mentioning the risk of nuclear war is the Cold War.” (Participant #35)

It is important to note that the distribution of frequencies in each proficiency level is relative due to the fact that only a single classification was given to each response during coding. In other words, having a higher frequency in one area influences the frequencies in other areas. For example, because the subcategory of understanding causal relationships and protagonists’ intentions represented 44.7% of responses in the high-intermediate group, the other subcategories necessarily represented smaller frequencies. This explains why the 2-b

subcategory of “marginally appropriate use of pragmatic knowledge” consists of only 5% or fewer responses across the proficiency levels (0.5% of 95 responses in the high-intermediate level, 5% of 120 responses in intermediate level, and 2.1% of 70 responses in the low-intermediate level).

In order to address the issue of a single category absorbing variance, the researcher changed the content analysis score range from 0 or 1 to 1, 2, or 3 (as shown in Table 4.10), since there are three levels (high, medium, and low) of understanding within each major category. This allowed the researcher to measure each participant’s think aloud performance based on the continuous scale coded by the two raters, as opposed to a categorical scale. For example, both raters assigned to 3-a (coherent understanding of the text) the response of participant #2 to passage one; this provided the highest score of 3. Thus, the average score for passage one between the two raters was 3. Similarly, the raters assigned the participant’s passage two response to 1-a (understanding causal relationship/the protagonist’s intention), which provided another average score of 3. For the passage three response, the raters had a slight disagreement. One rater coded it as 1-a (understanding causal relationship/the protagonist’s intention), whereas the other rater coded it as 1-b (superficial understanding of causal relationship/the protagonist’s intention). As 1-a is worth 3 points and 1-b is worth 2 points, the participant received an average of 2.5 points. The raters also had a disagreement on the response to passage four (3-b or moderate coherence vs. 2-b or marginal use of pragmatic knowledge). Regardless of the disagreement, the subject received an average score of 2 due to both subcategories being worth 2 points each. Regarding the passage five response, both raters regarded it as 3-b, which was worth another 2 points. As a total score, participant #2 received 12.5 points out of 15 possible points. The remaining subjects’ think aloud scores were averaged likewise.

Table 4.10 Sample Items for Content Analysis Scoring of Think Aloud Responses

			Overarching Ideas			Pragmatics			Coherence		
Passage #	Average Content Score	Rater	1a	1b	1c	2a	2b	2c	3a	3b	3c
1	3	M	0	0	0	0	0	0	1	0	0
1		A	0	0	0	0	0	0	1	0	0
2	3	M	1	0	0	0	0	0	0	0	0
2		A	1	0	0	0	0	0	0	0	0
3	2.5	M	1	0	0	0	0	0	0	0	0
3		A	0	1	0	0	0	0	0	0	0
4	2	M	0	0	0	0	1	0	0	0	0
4		A	0	0	0	0	0	0	0	1	0
5	2	M	0	0	0	0	0	0	0	1	0
5		A	0	0	0	0	0	0	0	1	0

With 1 (lowest) to 3 (highest) as the possible score range per item, the possible score range per subject was 5 (lowest) to 15 (highest), given that there were five passages. Therefore, the quality of each response generated by 57 participants in the experimental group was measured using the one to three scale based on the average scoring between the two raters. The actual think aloud scores per subject ranged from 7 (the lowest obtained by a low-intermediate-level participant) to 14 (the highest obtained by a high-intermediate-level participant). To determine whether any correlation existed between these scores and the participants' L2

proficiency, as well as the strength of the relationship, the researcher incorporated into the content analysis Pearson's r coefficient of correlation using think aloud scores and raw EPT scores.

Table 4.11 shows that there was a moderate positive correlation between the two variables (0.43321, $p < 0.05$). Since the correlation coefficient “ranges from -1.00 for a perfect negative correlation, through 0.00 for no relationship, to +1.00 for a perfect positive correlation,” the results indicated that the think aloud scores and EPT scores were moderately correlated (Goodwin, 2010, p. 330). The higher the participants' EPT scores were, the better they performed on the think aloud task, or vice versa, providing additional evidence for the linguistic threshold hypothesis.

Table 4.11 Descriptive Statistics and Pearson's Correlation Coefficient

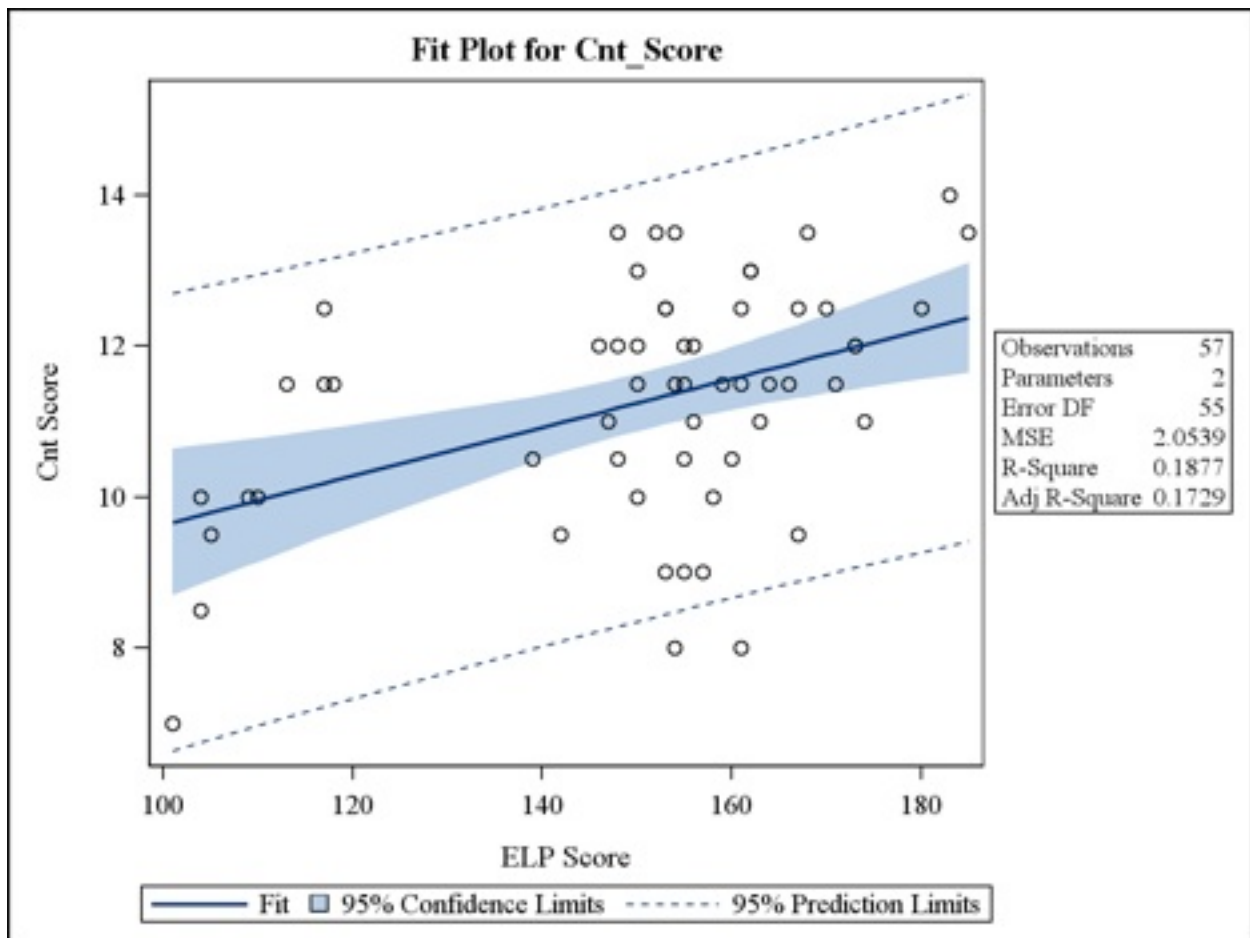
Variable	N	Mean	SD	Median	Min	Max
EPT Score	57	150.28070	21.18604	155.00	101.00	185.00
Cnt Score	57	11.25439	1.57583	11.500	7.00	14.00

Pearson's r (N = 57 Prob >/r/ under H0: Rho = 0)	EPT Score	Cnt Score
EPT Score	1.00	0.43321 0.0008
Cnt Score	0.43321 0.0008	1.00

This moderate correlation is represented in the scatterplot depicted in Figure 4.4. The data points cover a wide range of L2 fluency, which validate the diverse sample taken in this study. Although there are two separate clusters as opposed to one, they roughly fall in the same pattern. Ten out of 57 participants in the experimental condition are clustered at the bottom of

the scatterplot away from the remaining 47, which creates discontinuity. However, basically the same relationship is reflected for those clustered in the larger group. It is also important to note that such discontinuity is related only to the EPT scores, not the content/think aloud scores, suggesting that the gap may be due to sampling. In other words, there was a certain proficiency range (EPT scores ranging from 119 to 140) for which this study did not have any participants. Regardless of the gap between the two sets of clusters, the trend of positively correlated scores on the two measures is still the same.

Figure 4.4 Correlation between EPT Scores and Content Scores



What do these categorical data, which have been quantitatively represented, tell us about the quality of the participants' thinking aloud? What is a realistic description of their thought

processes? To answer these questions, the researcher qualitatively examined what participants reported in the think aloud protocol designed for this study. In the following section, the researcher discusses how participants across L2 proficiency levels described their thoughts and interpretations of the events elicited by the adjunct questions within the context of the three major categories (overarching ideas, pragmatic knowledge, and coherence).

Overarching ideas

As stated previously, the most frequently coded subcategory was understanding the causal relationships/a protagonist's intention under the category of overarching ideas. Consider the following response to the adjunct question, What are the effects of consuming high doses of vitamins?

(1) There are a [sic] bad effects in consuming high doses of vitamins. One of this [sic] is a liver damage, another one could be headaches, etc. For a healthy life, people must have a good and normal diet. (Participant #45)

This response was provided by a high-intermediate participant who summarized the causal relationship between the consumption of high doses of vitamins and health issues in his/her own words. It appeared that this participant was able to combine two different ideas embedded in the text as underlined in the text below:

Vitamin A, D, E, and K dissolve in fat, but not in water. They are stored in fat cells where they are available on demand. It also means that these vitamins can build up to toxic (poisonous or deadly) levels if taken well beyond normal requirements. For example, high doses of vitamin A can result in fatigue, headache, dizziness, blurred vision, dry skin, nausea, and liver damage. Vitamin D toxically occurs at just four to five times its

RDA (recommended dietary allowance), making vitamin D the most potentially toxic vitamin. Such high levels of the vitamin are reached using vitamin supplements, not through a normal diet. Cardiac (heart) and kidney damage can result (Yaworski, 2006).

In addition to understanding the cause (taking an excessive amount of vitamins) and the effect (health hazard), the reader successfully put the key information presented early in the passage (vitamins can build up to toxic levels if taken well beyond normal requirements) together with other key information found later in the passage (such high levels of the vitamin are reached using vitamin supplements, not through a normal diet). The idea that the danger of taking high doses of vitamins is irrelevant as long as vitamins are consumed through a normal diet was clearly comprehended. Since the reader's use of the term "a normal diet" came directly from the passage, s/he did not necessarily make an inference in connecting the two ideas. However, it is evident that sufficient working memory filled with the earlier information was involved in his/her comprehension of the information presented later in the text. This process is very similar to the way skilled readers make connections between different parts of a text in order to maintain a coherent understanding through the generation of situation models (Graesser et al., 1994; Kintsch, 1998/2007). Thus, this response (which will be referred to as 1) captures a trait of successful comprehension based on linking the incoming information with the reader's linguistic and world knowledge.

In contrast, neither of the following responses provided by an intermediate-level participant (2) and a low-intermediate-level participant (3) represents this rather sophisticated level of comprehension.

(2) I think that the effects of consuming high doses of vitamins can produce a lots [sic] of diseases and terrible damage in your liver, for example, like vitamin D is the most potentially toxic vitamin. (Participant #49)

(3) To consume high doses of vitamins can cause damage in the body organs. (Participant #120)

While these responses illustrate the participants' understanding of the effect of high doses of vitamins on one's health, they provide no evidence of connections being made between the two ideas across the passage, as was reflected in the response of the high-intermediate counterpart. Given the quality of each response, (1) received the highest think aloud score of 3 points, (2) received 2.5 points, and (3) received 1 point, which is the lowest score for this particular item. The participants' EPT scores were reported as 161, 154, and 139, respectively.

Pragmatic knowledge

Out of the three major categories, pragmatic knowledge was the least frequently reflected in responses by participants across proficiency levels. Observe the following responses to the adjunct question, What was Wilkie's intention in telling his audience that if they reelected Roosevelt, they would expect war in 1941? Since this Presidential candidate's intention is not explicitly stated in the text, the readers needed to make inferences in their responses.

(4) He *wanted to convince to [sic] the audience* that if they reelected Roosevelt, the USA would participate into the war. (Participant #86)

(5) It was [sic] ploy of election. So they were competing and Wilkie didn't want people to reelect Roosevelt by saying [sic] with him [sic] US will take [sic] into the war. (Participant #129)

A high-intermediate-level participant reported (4) and an intermediate-level participant reported (5) while reading the passage below:

In the 1940 Presidential election, challenger Wendell Wilkie charged that Franklin D. Roosevelt wanted to take the United States into the war (World War II). “If you reelect him,” he told one audience, “you may expect war in April 1941,” to which Roosevelt replied, “I have said this before, but I shall say it again and again and again: Your boys are not going to be sent into any foreign wars.” In November Roosevelt won the election with the electoral count 449 to 82 (Yaworski, 2006, p. 282).

Because the adjunct question is inferential, the intention of the protagonist Wilkie can be understood fully only when readers integrate the text information with their pragmatic knowledge of the world (e.g., what Presidential candidates often do during their campaign). In providing (4), the high-intermediate participant showed her/his pragmatic knowledge by using terms such as “convince the audience,” which is a standard practice politicians engaged in so that they can persuade voters to elect them instead of their opponents. The intermediate-level participant also showed her/his pragmatic knowledge in the first sentence in (5) by interpreting Wilkie’s intention as a “ploy of election.” As the participant describes, Presidential candidates are likely to strategically state something that is not necessarily true in order to gain some kind of advantage. This knowledge was well connected with the second sentence, “So they were competing, and Wilkie didn’t want people to reelect Roosevelt by saying [sic] with him [sic] US will take [sic] into the war.” In other words, according to this interpretation, Wilkie was playing a manipulative role, which is likely to happen in real political situations.

In contrast, a low-intermediate-level participant reported the rather non-pragmatic response that follows.

(6) Franklin D. Roosevelt wanted to take the U.S. into the (World War II). If he is elected, he said the boys are not going to go [sic] the war. (Participant #91)

Unlike (4) and (5), pragmatic knowledge in (6) is almost non-existent. In fact, in the first sentence this participant represented inaccurate knowledge of the world (Roosevelt wanting to take the country into the war), which is not consistent with Roosevelt's campaign promise explicitly stated in the passage. There is also inconsistency between the reader's first and second sentences. Besides not having sufficient pragmatic knowledge, this could be due to the reader's lack of basic comprehension of the text or a combination of the two. Responses such as (4), (5), and (6) support the idea that there is a linguistic threshold that enables L2 readers with sufficient L2 proficiency to appropriately use their domain knowledge (Barry & Lazarte, 1995, 1998, 2000; McNeil, 2010). Given the quality of each of the above responses, the three readers scored 2 points, 3 points, and 1.5 points, respectively, on this particular think aloud task.

Coherence

Similar to the pattern in the first subcategory, more high-intermediate participants represented their coherent understandings of the texts than lower-proficiency comprehenders. Of the 41 responses across L2 levels that were coded 3-a, the highest frequency count of 19 (46.3% of total responses in this subcategory) occurred among the high-intermediate participants, followed by the 11 responses (26.8%) of intermediate participants and the 11 responses (26.8%) of low-intermediate participants. There was no difference between the intermediate and low-intermediate levels in terms of how frequently participants demonstrated coherent understandings in their responses to the adjunct questions. This reveals that most coherent understandings of the passages resulted among students who had the highest level of English proficiency.

Consider the following response generated by one of the high-intermediate participants.

(7) Landfill linings are used. So, they create oxygen-free conditions to slow down bacteria and fungi. *For these reasons*, newspapers and hotdogs were not destroyed.

(Participant #58)

While missing the ultimate purpose of landfill linings (i.e., to prevent waste from leaking into the surrounding ground), this participant provided a coherent response to the adjunct question, What is the purpose of landfill linings? By paraphrasing one of the central details (i.e., create oxygen-free conditions that slow down bacterial and fungal action), the response maintained the key structure of the following text.

Modern waste disposal landfills (trash dumps) are covered and lined to prevent waste from leaking into the surrounding ground. Landfill linings and coverings create oxygen-free conditions that slow down bacterial and fungal action. Recent excavations (digging out) of old landfills have found 37-year-old newspapers that are still readable and five-year-old hotdogs; while hardly edible (not able to be eaten), they are at least recognizable (Yaworski, 2006).

In addition, the participant effectively used a transitional expression (*for these reasons*), which provided the response with coherence or a logical connection. In other words, the cause (landfill linings creating oxygen-free conditions to slow down bacterial and fungal actions) is well connected with its effect (waste not being decomposed). It appears that explaining the causality and intentionality by answering the adjunct question allowed this participant to construct a specific situation model, which facilitated his/her comprehension.

The following example by an intermediate-level participant also consists of a similar paraphrasing, yet the ending is rather disconnected from the rest of the response.

(8) The main purpose of landfill linings is creating oxygen-free condition which slow [sic] down bacterial and fungal action *because* old landfills are not working and still takes [sic] a long time to get rid of different trash dumps. (Participant #33)

Like the high-intermediate participant, this participant used a transition word (because) in an attempt to connect two different ideas. However, its following clause (old landfills are not working and still takes a long time to get rid of different trash dumps) does not have a causal relationship with its preceding clause (the main purpose of landfill linings is creating oxygen-free condition which slow [sic] down bacterial and fungal action). The comprehender did not quite build coherence, probably due to an insufficient understanding of the text and/or how this particular connective functions in a complex sentence (as shown in Goldman and Murray's [1992] study).

Compared to high-intermediate and intermediate participants, low-intermediate participants were more likely to repeat exactly what was stated in the text, as opposed to paraphrasing. Observe the following example.

(9) The purpose of landfill linings are [sic] *create oxygen-free conditions that slow down bacterial and fungal action.* (Participant #125)

While this participant identified the secondary purpose of the landfill linings, out of the 16 words in the sentence, 10 words in a row were taken directly from the text. This might have led to text-

boundedness, or a readers' over-reliance on the text information without making cognitive connections between different parts of the text.

Even when paraphrasing took place in other participants' responses, such as below, low-intermediate-level participants tended to focus on secondary details.

(10) The purpose of landfill linings is to reserve or keep something that can be recognize [sic] after a long time. (Participant #108)

Obviously, preserving objects such as newspapers and hotdogs for an extended period of time is not the primary purpose of landfill linings. It appeared that this participant did not quite comprehend that the 37-year-old newspapers and five-year-old hotdogs were found as a result of creating oxygen-free conditions in which bacterial and fungal action had been reduced. Instead, the participant focused on a less-salient idea, which is consistent with an earlier claim about struggling readers' inability to suppress irrelevant details (Gernsbacher, 1993; Gernsbacher & Robertson, 1995; Long et al., 2006).

Based on the findings from both the quantitative and qualitative analyses, the implications of the current study are discussed in the final chapter. The researcher will explore the relationship between participants' performance and their L2 proficiency, particularly from an instructional point of view. The chapter will conclude with suggestions for further research in L2 reading comprehension.

Chapter 5 - Discussion, Implications, and Conclusions

As the primary purpose of this study was to explore ways in which teachers can increase their L2 learners' reading comprehension through constructing situation models, this chapter provides an overall summary of the findings and their theoretical and practical significance.

Based on the implications of the study highlighted, the author will also discuss a set of recommendations for further research.

Summary of the Findings

The data analysis of this quasi-experimental study was conducted at three levels to address the following research questions: (1) To what extent is ELL students' ability to make L2 inferences affected by their L2 proficiency, (2) Can engaging ELL students in actively explaining target situation models (e.g., why and how a certain event occurred in the text as well as what the event caused) improve their L2 comprehension?, and (3) Does the effect of the intervention depend on L2 proficiency? Using descriptive and inferential analyses, the researcher tested four hypotheses:

- (1) Factual/non-inferential questions are more likely to be answered appropriately by L2 learners than those that require pragmatic inferences based on the reader's situation models (The main effect of the question type on comprehension)
- (2) Pragmatic inferences are more likely to be made by the learners with higher proficiency than those with lower proficiency, and the factual/non-inferential questions will be correctly answered at equal rates by learners of all levels of L2 proficiency. (The interaction between the question type and L2 proficiency on comprehension)
- (3) As ELL students actively engage in responding to explanatory/adjunct questions, they will better create situation models, and thus better comprehend the text. (The main effect of the intervention on comprehension)
- (4) The effectiveness of having ELL students respond to explanatory/adjunct questions will differ according to proficiency level. (The interaction between the intervention condition and L2 proficiency on comprehension)

Through the inferential statistical analysis, the researcher found that the first hypothesis was consistent with the data collected. As discussed in Chapter Four, there was an overwhelming effect of question type, such that all participants responded more accurately to the non-inferential questions than the inferential questions, which might seem a given. However, it was worthwhile to confirm the results of earlier L2 reading studies (Hammadou, 1991; Horiba, 1996; Kembo, 2001; Rai et al., 2011; Shimizu, 2005) that were reviewed in Chapter Two. Parallel to these scholars' findings, the current study has indicated that processing information at a pragmatic inference/situation model level is more difficult for all learners, as it requires the highest level of text representation (beyond surface and propositional representations).

Regarding the second hypothesis, although the quantitative data did not quite show a statistically significant interaction between question type and proficiency, there was a tantalizing effect of the question type interacting with L2 proficiency. As discussed in Chapter Four, there was almost an interaction between the within-group variable and the between-group variable. This tendency was more apparent with the high-proficiency-level participants than with their lower-proficiency counterparts. Thus, regardless of its statistical non-significance, the data trended in the direction that the second hypothesis predicted (i.e., pragmatic inferences are more likely to be made by the learners with higher proficiency than those with lower proficiency). In other words, lower-proficient ELL students are less likely to make appropriate pragmatic inferences during reading in their L2 because pragmatic inferences occur at the situation model level. Higher-proficiency learners respond to both inferential and non-inferential questions more appropriately than lower-proficiency learners.

Since there was no statistically significant main effect of the between-subject variable (experimental and control conditions) on accuracy, the third hypothesis was not supported by this

study. This means that eliciting the creation of situation models through asking and answering adjunct questions did not make a significant difference in increasing comprehension among participants in the experimental group. There are two reasons that might explain the non-significant effect of the intervention.

First, the adjunct question manipulation to promote the creation of situation models during reading (explaining the underlying situations embedded in the text vs. rereading the text) might not have been strong enough to make a difference between the two conditions. The control group might have benefited as much from reading the text twice as the experimental group benefitted from answering the adjunct questions regarding causality and intentionality. Thus, further research with a stronger manipulation is necessary to validate the importance of creating situation models; this point will be discussed further in the section for recommendations for future research.

Second, the effect of adjunct questions may vary from comprehender to comprehender. McMaster and colleagues (2012) conducted a noteworthy comparison between two types of struggling readers, paraphrasers and elaborators. In the study, the paraphrasers who repeated what was stated in the text were more likely to benefit from questions that promoted general inferences rather than causal inferences. In contrast, the elaborators who made inferences that were mostly unrelated to the central idea of the text tended to benefit from causal questions. Thus, the researchers concluded that asking questions to foster readers' understanding of causal relationships is beneficial to those who normally fail to generate inferences that are related to the central structures of the text. Since this paraphraser-elaborator dichotomy as a between-subject factor was not part of the current study's design, there was a chance that the participants were predominantly paraphrasers by nature who were less likely to benefit from the adjunct questions

presented. Given that three of the five questions were formulated specifically to elicit causal relationships, the comprehender quality might have had a positive impact on participants' performance.

As for the fourth hypothesis, other than the statistical significance regarding the effect of L2 proficiency on overall comprehension accuracy, no significant interaction between the experimental/control condition and L2 proficiency was observed. In other words, the effect of actively encouraging L2 readers to answer adjunct questions while reading did not depend on comprehenders' L2 proficiency. However, as discussed in Chapter Four, there was a trend of comprehension accuracy as a function of condition and L2 proficiency. The data trended in the direction of the hypothesis. The effectiveness of explanatory questions may well vary according to L2 proficiency, which would be consistent with the linguistic threshold hypothesis (Anderson, 1984; Carrell, 1991; Clarke, 1980).

The lack of interaction between question type and L2 proficiency (H2) as well as that between L2 proficiency and the experimental/control conditions (H4) could be partly due to the limited range of L2 proficiency represented by participants in the current study (i.e., high-intermediate being the highest level). As Table 4.4 in the previous chapter illustrates, the mean scores of the high-intermediate participants on comprehension questions were only 74 points for the experimental group and 69 points for the control group. If these participants' scores had been in the 80s or 90s, they might have shown the effect of question type and treatment. In other words, a lack of the main effect of higher proficiency participants in the current study (namely advanced level participants) might have contributed to the non-significant interaction between these variables. In order to validate this point, additional data from advanced level participants is

necessary. Such a scenario would allow the researcher to have greater variability among different proficiency levels, which might allow for a more effective testing of H2 and H4.

In sum, the quantitative measures revealed a significant question type effect (i.e., all participants answered non-inferential questions more appropriately than inferential questions). It also showed a non-significant trend in the direction consistent with the linguistic threshold hypothesis, in which participants' conscious actions of answering why and how questions led to better monitoring of their comprehension, but only for those with sufficient L2 proficiency.

The trend was consistent with the results of the qualitative measure, which was a variation of the think aloud protocol. As discussed in Chapter Four, the quality of participants' responses (i.e., content scores) was predicted by their L2 proficiency in terms of understanding causal and intentional relationships in the text. The higher participants' EPT scores, the better they responded to causality and intentionality-related questions. It is possible that the correlation between the EPT and content scores was due partly to the content scoring procedures used in this study. One may argue that the content scores were analyzed on the basis of L2 writing ability, which would not necessarily reflect participants' L2 comprehension, but rather their L2 production. Follow-up research could evaluate L2 readers' comprehension by having them respond in their L1 to avoid this issue.

On the other hand, this study has shown that having higher EPT scores was correlated with participants' greater ability to generate L2 inferences. However, correlation does not imply causation; as such, the author cannot conclude that there is a causal relationship between these variables (Goodwin, 2010). Nevertheless, content scores likely would increase as participants' ability to infer causal and intentional structures of the text increases. Naturally, the ability to draw inferences in the L2 increases together with L2 proficiency.

Linear relationships such as the positive correlation between the EPT scores and understanding of causal and intentional connections were not found in the other response categories involving pragmatic knowledge and coherent understanding. This was likely due to the way the analysis was designed. Namely, the raters assigned each response to only one of the three major categories, as opposed to all three categories. It appears that this particular response category, “understanding causal relationships and a protagonist’s intentions,” absorbed most of the variance in ratings of the students’ responses.

As an additional note on the L2 readers’ pragmatic knowledge, factors such as participants’ familiarity with the target culture and their age need to be taken into consideration, especially in interpreting their responses to passages with historical events. As discussed in Chapter Two, readers’ familiarity with the passage topic or content allows them to better predict and infer what the text is about (Afflerback, 1990; Garcia, Jimenez, & Pearson, 1998; Grabe, 2009; Graesser, Singer, & Trabasso, 1994; McNamara & Kintsch, 1996). It is safe to say that the more familiar participants are with US history, the better situation models they can create based on their pragmatic knowledge when reading passages 2 and 5. Given that both the 1940 US Presidential election (involving Roosevelt and Willkie) and the 1962 Cuban missile crisis took place more than five decades ago, neither of the events was part of the life experience of the participants, whose average age was 32. Furthermore, considering the participants’ nationality, they were very unlikely to make use of experiential knowledge to improve their comprehension of these passages. Instead, they would have had to rely on either their knowledge of US history, or more likely, their general domain knowledge (e.g., how politicians behave during their campaigns and/or how leaders deal with a national crisis) to make pragmatic inferences.

Probably it is best to avoid using reading passages that are strongly tied to US history and culture.

Given the findings of the quantitative and qualitative analyses, the author next will discuss both the theoretical and practical significance of the study.

Theoretical and Practical Implications of the Study

There are both theoretical and practical implications of the current study. Its theoretical significance relates to the non-significant, but suggestive trend that L2 readers who received the treatment in which they responded to adjunct questions were slightly more likely to create appropriate situation models if they had a certain level of proficiency in English. Specifically, the t-test comparing the experimental and control conditions at the high-intermediate level ($p = .218$, as opposed to $p = .698$ at the intermediate level and $p = .942$ at the low intermediate level) suggest that there might be something starting to happen at higher proficiency levels. This should be tested in follow-up research with an even higher proficiency level such as the advanced level. The practical significance of this study is connected to the types of questions (e.g., inferential questions related to causality and intentionality) that could potentially benefit L2 readers, specifically at higher proficiency levels. It is also concerned with the exploration of ways in which educators may be able to increase struggling comprehenders' ability to understand causality and intentionality while reading, though further research will be needed to produce more conclusive results on this issue.

At the theoretical level, this experimental study provided suggestive evidence for the linguistic threshold hypothesis in relation to L2 learners' ability to create appropriate situation models while reading. The linguistic threshold hypothesis was reviewed and discussed in Chapters Two and Four. The tantalizing trends found in the current study suggested that there

was a bigger effect of the question type and possibly also the experimental condition among the L2 participants with higher proficiency than among those who had lower proficiency in English. The data suggest that the influence of adjunct questions in creating situation models was slightly larger when the readers were at the highest L2 proficiency level (ranging between 161 and 191 on the EPT test). Thus, the contribution of these findings is to suggest a potential method of engaging L2 students with the text by using targeted adjunct questions (i.e., eliciting readers' explanations through questioning and self-monitoring of comprehension), which may improve L2 comprehension by facilitating the construction of more appropriate situation models. However, this will need to be tested in further research to draw stronger conclusions. Nevertheless, as reviewed in Chapter Two, McNeil's 2010 study confirmed a stronger effect of questioning on readers' comprehension than the influence of their existing domain knowledge, but only for the readers who had intermediate-level or higher L2 proficiency.

The current study also confirmed Carrell's (1984) claim of L2 readers' need for having a certain level of linguistic knowledge in order to make use of their domain knowledge while comprehending text. Consequently, the lower-proficient readers in her study did not go beyond the textbase level of comprehension. Like Carrell's study, in the current study it was found that lower-proficiency readers were less likely to succeed in activating their pragmatic knowledge to infer critical meanings represented in the passages (e.g., as shown in participant response example 9 in Chapter Four). In addition, the study supported Barry and Lazarte's (1995) finding of intermediate proficiency being the threshold for L2 readers' ability to process complex language structures, which in turn allowed them to appropriately utilize their domain knowledge. Reflecting a very similar threshold, for participants in the current study to use their pragmatic

knowledge to increase their L2 comprehension, they needed to have at least intermediate proficiency in English.

At the practical level, if having intermediate proficiency enables L2 readers to make appropriate pragmatic inferences based on their domain knowledge, then having lower-proficiency students activate their prior knowledge alone is insufficient to improve their reading comprehension. While it is important to stress cognitive processing such as making prior knowledge connections in L2 reading, educators need to assist students to improve their overall proficiency in the target language. One way to do so is to increase their vocabulary and grammatical knowledge in English, especially focusing on terms and structures used often in academic English discourse. Because L2 readers can only create an appropriate situation model for the text by making inferences based on their L2 grammatical competence and underlying world knowledge it makes sense that they need to possess a certain level of linguistic knowledge in order to integrate the text information with their prior knowledge effectively.

For higher-proficiency L2 readers who are above the linguistic threshold, teachers should encourage them to construct meaning through a situation model, in which their cognitive representations of the events described in text constantly change as they continue to read (Zwaan & Radvansky, 1998). This will allow them to apply more effectively what they already know to understanding the information that the text provides, while also maintaining situational coherence. By creating situation models through questioning and explaining, readers may be able to better link incoming text information to discourse elements, including causality and intentionality (Magliano, Zwaan, & Graesser, 1999; Zwaan et al., 1995; Zwaan & Brown, 1996; Zwaan & Radvansky, 1998). As discussed in Chapter Two, this allows readers to monitor their comprehension by keeping track of why an event occurred and what the protagonist's goal was.

Nevertheless, these conclusions from the current study are tentative, due to the lack of statistical significance of the relevant hypotheses. Further research with a stronger adjunct questioning manipulation may be able to provide more conclusive results.

Recommendations for Future Research

Having examined the data from the current study in relation to the hypotheses, while also considering the above implications, the author proposes the following recommendations for future research on L2 reading research of this sort: (1) increase the strength of the adjunct question manipulation so that it has a greater likelihood of having an effect on L2 learners' reading comprehension, (2) collect data from participants with a wider range of L2 proficiency, and (3) pre-assess participants' comprehender quality (i.e., paraphrasers vs. elaboraters).

To increase the power of the adjunct question manipulation, it is likely necessary to increase the learners' engagement with the text at a meaningful level by adding a practice element to the design and having participants generate their own questions. As suggested by the qualitative analysis of the participants' answers to the adjunct questions, giving good answers to adjunct questions is a skill that takes time to develop. The design of the current study included a total of five passages, each having an adjunct question in the experimental condition, so that by the end the experimental participants were used to answering adjunct questions related to causality and intentionality. However, the current study lacked practice texts and adjunct questions prior to the actual assessment. Clearly, making an inference about what is not directly stated requires a rather sophisticated level of comprehension in one's L2. The participants in the experimental group might have benefitted from working with several practice texts to create the mindset of answering inferential adjunct questions related to causality and intentionality. In addition to the practice element, the participants could also talk about the underlining meaning of

the text and then generate their own questions regarding the underlying meaning of each passage. Such additional activities involving L2 learners in generating inferences while constructing coherent situation models would likely increase the power of the experimental manipulation. While engaging the treatment group in such ways, alternative tasks for the control condition need to be added (e.g., underlining what the participants think is important for each passage, as often occurred in the current study).

As discussed previously, it is possible that not including participants from the advanced-proficiency level in the current study might have contributed to the non-significant interactions between L2 proficiency on the one hand, and condition and question type on the other. Specifically, if the high-intermediate group had been proficient enough to have scored 80 points or higher on the comprehension task, there may have been a greater chance of finding the predicted interactions between L2 proficiency and condition, and L2 proficiency and question type, as suggested by the linguistic threshold hypothesis. Although the high-intermediate level was the highest proficiency group available to participate in this study, a suggestion for future research would be to collect data from a wider range of L2 proficiencies, namely from the low-intermediate to advanced levels, especially adding high students and/or native speakers. Due to the nature of the English language/professional development program in which the participants were involved (i.e., professional development for English teachers from Ecuador), however, inclusion of individuals with lower L2 proficiency than the low-intermediate level is not possible. Fortunately, however, the program has evolved since the collection of the current data to include a master's program component. University acceptance criteria for this program include a language assessment piece. To be eligible for participation, students have to obtain 200 points or higher on the EPT test, or 20 points or higher on each of the TOEFL subtests of

listening, writing, and reading. Accepted students will likely represent the advanced proficiency level needed for possible future research.

In addition, it may be useful to take into consideration participants' comprehender type for further research. The recent study by McMaster et al. (2012) distinguished between elaborators and paraphrasers, and found differences in the benefits to comprehension of different types of questions. As described previously, elaborators were defined as readers who were more likely to draw elaborative inferences from the text, whereas paraphrasers were defined as those who simply repeated the text content. Specifically, McMaster et al. (2012) showed that only elaborators (as opposed to paraphrasers) benefitted from causal questions. More specifically, when the elaborators failed to understand the central ideas of the text because they made inappropriate inferences, causal questions helped them draw more appropriate inferences. Paraphrasers on the other hand, benefited from questions that elicited general inferences, but not from the causal questions. Thus, it would be worthwhile to assess what type of comprehenders the future participants are prior to the actual data collection. This could be done by asking them to verbally express their interpretation of a given text in written form, then to be coded for scoring. Including such a comprehender type as a between-subjects variable in turn would provide an additional dimension when analyzing data in terms of the influence of a comprehender type on the effect of causal questions. It would also be interesting to see if such future research would replicate the above described findings of McMaster and colleagues (2012).

Conclusions

By asking purposeful questions that promote the creation of situation models among L2 learners, educators can engage students' in higher-level cognitive processing while reading, specifically, encouraging them to make connections between the text information and their

pragmatic knowledge. This process likely allows the students to draw appropriate inferences, while also continually revising their situation models in order to maintain coherence. The current study replicated previous studies in terms of the main effect of question type on comprehension, such that lower proficiency students (here, lower-intermediate level) were only able to correctly answer factual (non-inferential) questions, whereas higher proficiency students (here, upper-intermediate level) were also able to accurately answer pragmatic inference questions (Hammadou, 1991; Horiba, 1996; Kembo, 2001; Rai et al., 2011; Shimizu, 2005). This finding is consistent with the linguistic threshold hypothesis (Clarke, 1980; Anderson, 1984; Carrell, 1991; Lee & Schallert, 1997), which argues that in order for L2 readers to use their L1 reading skills to improve their L2 reading comprehension, they require a minimal level of L2 linguistic proficiency. The results of the current study also suggest several potential relationships among the following set of variables: the interaction between question type (IV) and L2 proficiency level (IV) on L2 reading comprehension accuracy (DV); the main effect of the adjunct question intervention (IV) on L2 reading comprehension accuracy (DV); and the interaction between the intervention (IV) and L2 proficiency (IV) on L2 reading comprehension accuracy (DV). As noted above, consistent with the linguistic threshold hypothesis (Clarke, 1980; Anderson, 1984; Carrell, 1991; Lee & Schallert, 1997), the higher proficiency readers made more appropriate inferences than lower proficiency readers. More importantly, however, consistent with the linguistic threshold hypothesis, there was a statistically non-significant trend such that the adjunct questions appeared to have a greater impact on the comprehension of higher proficiency L2 readers than their lower proficiency counterparts, and this should be tested in future research. If further research bears out this suggestive trend, such information could allow educators to scaffold their instruction in terms of gradually shifting from factual/non-inferential questions to

inferential questions as students' level of L2 comprehension increases. In this way, educators can assist L2 readers to make academic gains in L2 literacy by both focusing on increasing their overall L2 linguistic proficiency and monitoring their own L2 comprehension through think-aloud based questioning.

Through this research, the author has gained insights into the role of questioning in creating situation models. Specifically, the current study has shown the importance of continued interaction with the text throughout the reading process, while activating knowledge structures. It is crucial for educators to train their students to utilize adjunct questions as a guide to process the text content beyond the factual level and take it to the inferential level. This will allow the students to keep track of the situation models they are constructing and ensure that they are appropriate and coherent. Further research with a stronger adjunct question manipulation, a wider range of L2 proficiency, and an additional between-subjects variable (i.e., comprehender type) are promising directions to follow-up on the results of the current study. In addition, it is best to incorporate a differentiated instructional component (e.g., modified adjunct questions) into a future study to see if all students including lower proficiency students can read for meaning that is beyond the textbase level of comprehension. It would be informative to find what sort of inferential questions can benefit lower proficiency students and/or how a teacher can deliver instruction to reach a wide range of students with varied L2 proficiency.

References

- Afflerback, P. P. (1990). The influence of prior knowledge on expert readers' main idea construction strategies. *Reading Research Quarterly* 25, 31-46.
- Ajudeh, P. (2003). Schema theory-based pre-reading tasks: A neglected essential in the ESL reading class. *The Reading Matrix*, 3(1), 1-14.
- Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 1-24). London: Longman.
- Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 255-292). New York: Longman.
- Anderson, R. C., Reynolds, R. C., Schllert, D. L., & Goetz, E. T. (1977). Frameworks for comprehending discourse. *American Educational Research Journal*, 14, 367-381.
- Barry, S., & Lazarte, A. A. (1995). Embedded clause effects on recall: Does high prior knowledge of content domain overcome syntactic complexity in students of Spanish? *The Modern Language Journal*, 79(4), 491-504.
- Barry, S., & Lazarte, A. A. (1998). Evidence for mental models: How do prior knowledge, syntactic complexity, and reading topic affect inference generation in a recall task for nonnative readers of Spanish? *The Modern Language Journal*, 82(2), 176-193.
- Barry, S., & Lazarte, A. A. (2000, April). *A comparison of first and second language readers' situation model as evidence in a recall task*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*, 717-726.
- Brown, C. M. (1998). L2 reading: An update on relevant L1 research. *Foreign Language Annuals, 31*(2), 191-202.
- Cain, K., & Oakhill, J. (2007). *Children's comprehension problems in oral and written language*. New York: Guilford.
- Carrell, P. L. (1983). Three components of background knowledge in reading comprehension. *Language Learning, 33*(2), 183-207.
- Carrell, P. L. (1984). Evidence of a formal schema in second language comprehension. *Language Learning, 34*(2), 87-108.
- Carrell, P. L. (1987). Content and formal schemata. *TESOL Quarterly, 21*, 461-487.
- Carrell, P. L. (1991). Second language reading: Reading ability or language proficiency? *Applied Linguistics, 12*(2), 159-179.
- Carrell, P. L., Gajdusek, L., & Wise, T. (1998). Metacognition and EFL/ESL reading. *Instructional Science, 26*, 97-112.
- Carroll, D. W. (2008). *Psychology of language* (5th ed.). Belmont, CA: Thomson Wadsworth.
- Clarke, M. (1980). The short circuit hypothesis of ESL reading: Or when language competence interferes with reading performance. *Modern Language Journal, 64*, 203-209.
- Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. New York: Multilingual Matters.
- Field, A. (2009). *Discovering statistics using SPSS*. New York: SAGE.

- Folse, K. (2009). Why K-12 teachers need to know about ESL grammar issues. *TESOL Journal*. Retrieved August 19, 2013, from [http://www.tesol.org/read-and-publish/journals/other-serial-publications/compleat-links/compleat-links-volume-6-issue-3-4-\(october-2009\)/why-k-12-teachers-need-to-know-about-esl-grammar-issues](http://www.tesol.org/read-and-publish/journals/other-serial-publications/compleat-links/compleat-links-volume-6-issue-3-4-(october-2009)/why-k-12-teachers-need-to-know-about-esl-grammar-issues)
- Foltz, P. W. (2003). Quantitative cognitive models of text and discourse processing. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 487-523). Mahwah, NJ: Lawrence Erlbaum.
- Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with latent semantic analysis. *Discourse Processes*, 25, 285-307.
- Garcia, G. E., Jimenez, R. T., & Person, P. J. (1998). Metacognition, childhood bilingualism and reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 193-219). Mahwah, NJ: Lawrence Erlbaum.
- Gay, L. R., Millis, G. E., & Airasian, P. (2006). *Educational research: Competencies for analysis and applications*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science*, 4, 294-298.
- Gernsbacher, M. A., & Robertson, R. R. (1995). Reading skills and suppression revisited. *Psychological Science*, 6(3), 165-169.
- Goldenberg, C. (2008). Teaching English language learners: What the research does—and does not—say. *American Educator*, 32(2), 8-44.
- Goldman, S. R., & Murray, J. D. (1992). Knowledge of connectors as cohesion devices in text: A comparative study of native-English and English-as-a-second-language speakers. *Journal of Educational Psychology*, 84, 504-519.

- Goodwin, C. J. (2010). *Research in psychology: Methods and design*. NJ: Wiley.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge Applied Linguistics.
- Graesser, A. C. (2007). An introduction to strategic reading comprehension. In D. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 3-26). Mahwah, NJ: Lawrence Erlbaum.
- Graesser, A. C., McMahan, C. I., & Johnson, B. K. (1994). Question asking and answering. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 517-538). San Diego, CA: Academic Press.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371-395.
- Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, *48*, 163-189.
- Hakuta, K. (1990). Language and cognition in bilingual children. In A. Padilla, H. Fairchild, & C. Valadez (Eds.). *Bilingual education: Issues and strategies* (pp. 47-59). Newbury Park: Sage.
- Hammadou, J. (1991). Interrelationships among prior knowledge, inference, and language proficiency in foreign language reading. *The Modern Language Journal*, *75*(1), 27-38.
- Hansen, J. J. (1981). An inferential comprehension strategy for use with primary grade children. *The Reading Teacher*, *34*(6), 665-669.
- Horiba, Y. (1996). Comprehension processes in L2 reading: Language competence, textual coherence, and inferences. *Studies in Second Language Acquisition*, *18*(4), 433-473.

- Horiba, Y., van den Broek, P. W., & Fletcher, C. R. (1993). Second language readers' memory for narrative texts: Evidence for structure-preserving top-down processing. *Language Learning, 43*(3), 345-372.
- Kembo, J. A. (2001). Testing of inferencing behaviour in a second language. *International Journal of Bilingual Education and Bilingualism, 4*(2), 77-96.
- Kintsch, E. (2005). Comprehension theory as a guide for the design of thoughtful questions. *Topics in Language Disorders, 25*(1), 51-64.
- Kintsch, W. (2007). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford, UK: Pergamon Press.
- Lee, J., & Schallert, D. L. (1997). The relative contribution of L2 language proficiency and L1 reading ability to L2 reading performance: A test of the threshold hypothesis in an EFL context. *TESOL Quarterly, 31*(4), 713-739.
- Long, D. L., & Golding, J. M. (1993). Superordinate goal inferences: Are they automatically encoded during comprehension? *Discourse Processes, 16*, 55-73.
- Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction, 21*(3), 251-283.
- Magliano, J. P., Millis, K. K., Ozuru, Y., & Mcnamara, D. S. (2007). A multidimensional framework to evaluate reading assessment tools. In D. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 107-136). New York: Lawrence Erlbaum.

- Magliano, J. P., Trabasso, T., & Graesser, A. (1999). Strategic processing during comprehension. *Journal of Educational Psychology, 91*, 615-629.
- Magliano, J. P., Zwaan, R. A., & Graesser, A. (1999). The role of situational continuity in narrative understanding. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 219-245). Mahwah, NJ: Lawrence Erlbaum.
- Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology, 11*(2), 194-201.
- McKoon, G., & Ratcliff, R. (1992). Inferences during reading. *Psychological Review, 99*, 440-466.
- McMaster, K. L., van den Broek, P., Espin, C. A., White, M. J., Rapp, D. N., Kendeou, P., & Carlson, S. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences, 22*(1), 100-111.
- McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38*, 10-30.
- McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes, 22*, 247-288.
- McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51, pp. 297-384). Burlington, NY: Academic Press.
- McNeil, L. (2010). Investigating the contributions of background knowledge and reading comprehension strategies to L2 reading comprehension: An exploratory study. *Springer*

- Science and Business Media*. Retrieved March 28, 2011, from <http://dx.doi.org/10.1007/s11145-010-9230-6>.
- Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes, 26*, 131-157.
- Nassaji, H. (2007). Schema theory and knowledge-based reading comprehension: A need for alternative perspectives. *Language Learning, 57*, 79-113.
- Oakhill, J., & Cain, K. (2007). Issues of causality in children's reading comprehension. In D. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 47-71). New York: Lawrence Erlbaum.
- Oakhill, J., & Yuill, N. (1996). Higher order factors in comprehension disability: Process and remediation. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties* (pp. 69-92). Mahwah, NJ: Lawrence Erlbaum.
- Pressley, M., & Harris, K.R. (1994). Increasing the quality of educational intervention research. *Educational Psychology Review, 6*(3), 191-208.
- Rai, M. K., Loschky, L. C., Harris, R. J., Peck, N. R., & Cook, L. G. (2011). The effects of stress and working memory on foreign language readers' inferential processing during comprehension. *Language Learning, 61*(1), 187-218.
- Schwartz, A. I., & Kroll, J. F. (2006). Language comprehension in bilingual speakers. In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 967-999).
- Shimizu, M. (2005). *Inference generation processes of Japanese EFL learners: Effects of questioning on their reading comprehension*. (Unpublished doctoral dissertation). University of Tsukuba, Japan.

- Singer, M. (1994). Discourse inference processes. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 479-515). San Diego, CA: Academic Press.
- Smith, F. (1988). *Understanding reading: A psycholinguistic analysis of reading and learning to read* (4th ed.). Mahwah, NJ: Lawrence Erlbaum.
- Steffensen, M. S., Joag-dev, C., & Anderson, R. C. (1979). A cross-cultural perspective on reading comprehension. *Reading Research Quarterly*, *15*(1), 10-29.
- St. George, M., Mannes, S., & Hoffman, J. (1997). Individual difference in inference generation: An ERP analysis. *Journal of Cognitive Neuroscience*, *9*(6), 776-787.
- Trabasso, T., & Magliano, J. P. (1996). How do children understand what they read and what can we do to help them? In M. Graves, P. van den Broek, & B. Taylor (Eds.), *The first R: every child's right to read* (pp. 160-188). New York: Teachers College Press.
- Uso-Juan, E. (2006). The compensatory nature of discipline-related knowledge and English-language proficiency in reading English for academic purposes. *The Modern Language Journal*, *90*(2), 210-227.
- van den Broek, P. (1994). Comprehension and memory for narrative texts: Inferences and coherence. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539-88). San Diego, CA: Academic Press.
- van den Broek, P., & Espin, C. A. (2012). Connecting cognitive theory and assessment: Measuring individual differences in reading comprehension. *School Psychology Review*, *41*, 315-325.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies in discourse comprehension*. New York: Academic Press.

- Whitney, P., Ritchie, B. G., Clark, M. B. (1991). Working memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes, 14*, 133-145.
- Wragg, E. C. (2012). *An introduction to classroom research*. New York: Routledge.
- Yamashita, J. (2002). Mutual compensation between L1 reading ability and L2 language proficiency in L2 reading comprehension. *Journal of Research in Reading, 25*(1), 81-95.
- Yaworski, J. (2006). *Getting ahead: Fundamentals of college reading*. New York: Pearson.
- Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes, 21*, 289-327.
- Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6*, 292-297.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162-185.
- Zwaan, R. A., & Rapp, D. (2006). Discourse comprehension. In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 725-64). Burlington, MA: Academic Press.
- Zwaan, R. A., & Singer, M. (2003). Text comprehension. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 83-121). Mahwah, NJ: Lawrence Erlbaum.

Appendix A - Inference Passages for the Control Condition

Name: _____

Directions: Please read the following passage twice and then answer the questions below.

The two ladies ran to meet the bus with four shopping bags in each hand. It seemed as though all of the dogs in New York City had been following them very closely. The driver closed the door just in time to keep the dogs out of the bus. As they sat down, they began to recount their purchases – five loaves of bread, two hams, and twelve rings of garlic sausage. What a feast they would have during the Ukrainian Easter holidays!

Q1: Why did ladies recount their purchases? Select the best answer by circling its number.

Because:

- (1) Because the ladies were suspicious about the shopkeeper.
- (2) Because the ladies thought that they purchased two hams instead of twelve rings of sausages.
- (3) Because the ladies thought that the dogs might have gotten the meat.

Q2: Why did the driver close the door? Select the best answer by circling its number.

Because:

- (1) The driver wanted to keep the dogs out of the bus.
- (2) The driver was running late.
- (3) The driver didn't want all of the dogs in New York City to follow the bus.

Q3: What is the ladies' cultural heritage suggested in the passage? Select the best answer by circling its number.

- (1) No cultural heritage suggested in the passage.
- (2) The ladies are of Ukrainian ancestry.
- (3) The ladies are of Italian ancestry.

Name: _____

Directions: Please read the following passage twice and then answer the questions below.

In the 1940 Presidential election, challenger Wendell Wilkie charged that Franklin D. Roosevelt wanted to take the United States into the war (World War II). “If you reelect him,” he told one audience, “you may expect war in April 1941,” to which Roosevelt replied. “I have said this before, but I shall say it again and again and again: Your boys are not going to be sent into any foreign wars.” In November Roosevelt won the election with the electoral count 449 to 82.

Q1: Why did Wilkie argue that people should vote for him instead of Roosevelt? Select the best answer by circling its number.

- (1) Wilkie wanted to help Roosevelt get reelected because Roosevelt promised to lead the country into the war.
- (2) Wilkie told one audience that Roosevelt would lead the country into war because he was trying to convince Roosevelt to be involved in the war.
- (3) Wilkie tried to convince American people that Roosevelt would go to the war if they elected Roosevelt.

Q2: According to the author, what position did Wilkie take about the US entering into the war? Select the best answer by circling its number.

- (1) Wilkie thought that US should enter into the war.
- (2) Wilkie opposed the US entry into the war.
- (3) Wilkie believed that a war-time president would be popular.

Q3: What was the major issue in the 1940 Presidential campaign? Select the best answer by circling its number.

- (1) The major issue in the 1940 Presidential election was which candidate would be most likely to keep the US out of foreign wars.
- (2) The major issue in the 1940 Presidential election was whether or not the US would enter foreign wars if Roosevelt were elected.
- (3) The major issue in the 1940 Presidential election was whether or not Wilkie would keep the US out of foreign wars.

Name: _____

Directions: Please read the following passage twice and then answer the questions below.

Vitamin A, D, E, and K dissolve in fat, but not in water. They are stored in fat cells where they are available on demand. It also means that these vitamins can build up to toxic (poisonous or deadly) levels if taken well beyond normal requirements. For example, high doses of vitamin A can result in fatigue, headache, dizziness, blurred vision, dry skin, nausea, and liver damage. Vitamin D toxically occurs at just four to five times its RDA (recommended dietary allowance), making vitamin D the most potentially toxic vitamin. Such high levels of the vitamin are reached using vitamin supplements, not through a normal diet. Cardiac (heart) and kidney damage can result.

Q1: According to the author, what is the causal relationship between high doses of vitamins and health? Select the best answer by circling its number.

- (1) High doses of vitamins can cause various health issues, including fatigue, headache, dizziness, blurred vision, dry skin, nausea, liver, kidney, and cardiac damage.
- (2) Vitamin A is the most toxic vitamin at lowered doses.
- (3) People can easily get toxic levels of vitamins by eating too much of some foods.

Q2: What does the author imply by saying, “vitamins can build up to toxic levels if taken well beyond normal requirements”? Select the best answer by circling its number.

- (1) Fat soluble vitamin supplements should not be taken daily because they can build vitamins up to toxic levels.
- (2) High doses of vitamin D can be safe because its toxicity occurs at just four to five times its RDA.
- (3) A normal diet can cause cardiac and kidney damage because vitamin D is stored in fat cells.

Q3: What does the author believe about dangerous doses of vitamin D? Select the best answer by circling its number.

- (1) We don't need to worry about vitamin D unless it is greater than four to five times its RDA.
- (2) We don't need to worry about dangerous doses of vitamin D if we get them entirely through food.
- (3) We need to worry about vitamin D only when it dissolves in fat cells.

Name: _____

Directions: Please read the following passage twice and then answer the questions below.

Modern waste disposal landfills (trash dumps) are covered and lined to prevent waste from leaking into the surrounding ground. Landfill linings and coverings create oxygen-free conditions that slow down bacterial and fungal action. Recent excavations (digging out) of old landfills have found 37-year-old newspapers that are still readable and five-year-old hotdogs; while hardly edible (not able to be eaten), they are at least recognizable.

Q1: What causes modern waste disposal facilities to slow down bacterial and fungal action? Select the best answer by circling its number.

- (1) Modern facilities are made of leakage-free materials so that waste cannot escape into the surrounding ground.
- (2) Modern facilities use landfill linings that hinder bacteria and fungi from growing.
- (3) Modern facilities produce less oxygen so that they can increase bacterial and fungal action.

Q2: Why were the 37-year-old newspapers readable and five-year-old hotdogs still recognizable in old landfills? Select the best answer by circling its number.

- (1) Because the newspapers and hotdogs were made of substances that never break down.
- (2) Because of the oxygen-free conditions that the landfill linings create, newspaper and hotdogs never rot.
- (3) Because many substances, including newspapers and hotdogs, break down very slowly when covered with landfill linings.

Q3: What makes modern landfills less dangerous to the environment than older landfills? Select the best answer by circling its number.

- (1) The covering and linings make modern landfills less dangerous to water supplies than older landfills.
- (2) There are fewer bacteria and fungi to make materials decompose in modern landfills than in older landfills.
- (3) Recent excavations have found that materials do not decompose as rapidly in older landfills as they do in modern landfills.

Name: _____

Directions: Please read the following passage twice and then answer the questions below.

The Cuban missile crisis of 1962 was the most terrifying confrontation of the Cold War. American aerial photographs taken revealed that the USSR had begun to place what US President John F. Kennedy considered offensive missiles (missiles designed to attack another country) on Cuban soil. Kennedy went on nationwide TV to tell the American people about the missiles and to demand their removal. He declared that the United States would not shrink from the risk of nuclear war and announced a naval “quarantine” around Cuba to prevent Soviet Union ships from bringing in additional missiles.

Q1: What did Kennedy mean by “the United States would not shrink from the risk of nuclear war”? Select the best answer by circling its number.

- (1) Kennedy did not want to risk the lives of people in the United States over Cuba.
- (2) Kennedy would negotiate to prevent a nuclear war in any way he could.
- (3) Kennedy would risk a nuclear war for the removal of Cuban missiles.

Q2: Why did the author refer to the Cuban missile crisis as “the most terrifying confrontation of the Cold War?” Select the best answer by circling its number.

- (1) Because it would increase the likelihood of nuclear war affecting the whole world.
- (2) Because it would increase the number of nuclear weapons set up by both United States and the Soviet Union.
- (3) Because it would decrease the chance of removing Cuban missiles.

Q3: What did the US speculate the Soviet Union was planning to do? Select the best answer by circling its number.

- (1) The Soviet Union was planning to point their nuclear missiles at Cuba.
- (2) The Soviet Union was planning to attack the US from Cuba.
- (3) The Soviet Union was planning to reduce the number of their nuclear weapons in Europe by bringing them to Cuba by ships.

Appendix B - Inference Passages for the Experimental Condition

Name: _____

Directions: Why were dogs following the ladies?

Try to give an explanation to the above question while reading the passage below. After reading only once, write your explanation in the blanks below the passage.

The two ladies ran to meet the bus with four shopping bags in each hand. It seemed as though all of the dogs in New York City had been following them very closely. The driver closed the door just in time to keep the dogs out of the bus. As they sat down, they began to recount their purchases – five loaves of bread, two hams, and twelve rings of garlic sausage. What a feast they would have during the Ukrainian Easter holidays!

Explanation: _____

Now answer questions 1-3 below.

Q1: Why did ladies recount their purchases? Select the best answer by circling its number.

Because:

- (1) Because the ladies were suspicious about the shopkeeper.
- (2) Because the ladies thought that they purchased two hams instead of twelve rings of sausages.
- (3) Because the ladies thought that the dogs might have gotten the meat.

Q2: Why did the driver close the door? Select the best answer by circling its number.

Because:

- (1) The driver wanted to keep the dogs out of the bus.
- (2) The driver was running late.
- (3) The driver didn't want all of the dogs in New York City to follow the bus.

Q3: What is the ladies' cultural heritage suggested in the passage? Select the best answer by circling its number.

- (1) No cultural heritage suggested in the passage.
- (2) The ladies are of Ukrainian ancestry.
- (3) The ladies are of Italian ancestry.

Name: _____

Directions: What was Wilkie's intention in telling his audience that if they reelected Roosevelt, they would expect war in 1941?

Try to give an explanation to the above question while reading the passage below. After reading only once, write your explanation in the blanks below the passage.

In the 1940 Presidential election, challenger Wendell Wilkie charged that Franklin D. Roosevelt wanted to take the United States into the war (World War II). "If you reelect him," he told one audience, "you may expect war in April 1941," to which Roosevelt replied. "I have said this before, but I shall say it again and again and again: Your boys are not going to be sent into any foreign wars." In November Roosevelt won the election with the electoral count 449 to 82.

Explanation: _____

Q1: Why did Wilkie argue that people should vote for him instead of Roosevelt? Select the best answer by circling its number.

- (1) Wilkie wanted to help Roosevelt get reelected because Roosevelt promised to lead the country into the war.
- (2) Wilkie told one audience that Roosevelt would lead the country into war because he was trying to convince Roosevelt to be involved in the war.
- (3) Wilkie tried to convince American people that Roosevelt would go to the war if they elected Roosevelt.

Q2: According to the author, what position did Wilkie take about the US entering into the war? Select the best answer by circling its number.

- (1) Wilkie thought that US should enter into the war.
- (2) Wilkie opposed the US entry into the war.
- (3) Wilkie believed that a war-time president would be popular.

Q3: What was the major issue in the 1940 Presidential campaign? Select the best answer by circling its number.

- (1) The major issue in the 1940 Presidential election was which candidate would be most likely to keep the US out of foreign wars.
- (2) The major issue in the 1940 Presidential election was whether or not the US would enter foreign wars if Roosevelt were elected.
- (3) The major issue in the 1940 Presidential election was whether or not Wilkie would keep the US out of foreign wars.

Name: _____

Directions: What are the effects of consuming high doses of vitamins?

Try to give an explanation to the above question while reading the passage below. After reading only once, write your explanation in the blanks below the passage.

Vitamin A, D, E, and K dissolve in fat, but not in water. They are stored in fat cells where they are available on demand. It also means that these vitamins can build up to toxic (poisonous or deadly) levels if taken well beyond normal requirements. For example, high doses of vitamin A can result in fatigue, headache, dizziness, blurred vision, dry skin, nausea, and liver damage.

Vitamin D toxically occurs at just four to five times its RDA (recommended dietary allowance), making vitamin D the most potentially toxic vitamin. Such high levels of the vitamin are reached using vitamin supplements, not through a normal diet. Cardiac (heart) and kidney damage can result.

Explanation: _____

Q1: According to the author, what is the causal relationship between high doses of vitamins and health? Select the best answer by circling its number.

- (1) High doses of vitamins can cause various health issues, including fatigue, headache, dizziness, blurred vision, dry skin, nausea, liver, kidney, and cardiac damage.
- (2) Vitamin A is the most toxic vitamin at lowered doses.
- (3) People can easily get toxic levels of vitamins by eating too much of some foods.

Q2: What does the author imply by saying, “vitamins can build up to toxic levels if taken well beyond normal requirements”? Select the best answer by circling its number.

- (1) Fat soluble vitamin supplements should not be taken daily because they can build vitamins up to toxic levels.
- (2) High doses of vitamin D can be safe because its toxicity occurs at just four to five times its RDA.
- (3) A normal diet can cause cardiac and kidney damage because vitamin D is stored in fat cells.

Q3: What does the author believe about dangerous doses of vitamin D? Select the best answer by circling its number.

- (1) We don't need to worry about vitamin D unless it is greater than four to five times its RDA.
- (2) We don't need to worry about dangerous doses of vitamin D if we get them entirely through food.
- (3) We need to worry about vitamin D only when it dissolves in fat cells.

Name: _____

Directions: What is the purpose of landfill linings?

Try to give an explanation to the above question while reading the passage below. After reading only once, write your explanation in the blanks below the passage.

Modern waste disposal landfills (trash dumps) are covered and lined to prevent waste from leaking into the surrounding ground. Landfill linings and coverings create oxygen-free conditions that slow down bacterial and fungal action. Recent excavations (digging out) of old landfills have found 37-year-old newspapers that are still readable and five-year-old hotdogs; while hardly edible (not able to be eaten), they are at least recognizable.

Explanation: _____

Q1: What causes modern waste disposal facilities to slow down bacterial and fungal action? Select the best answer by circling its number.

- (1) Modern facilities are made of leakage-free materials so that waste cannot escape into the surrounding ground.
- (2) Modern facilities use landfill linings that hinder bacteria and fungi from growing.
- (3) Modern facilities produce less oxygen so that they can increase bacterial and fungal action.

Q2: Why were the 37-year-old newspapers readable and five-year-old hotdogs still recognizable in old landfills? Select the best answer by circling its number.

- (1) Because the newspapers and hotdogs were made of substances that never break down.
- (2) Because of the oxygen-free conditions that the landfill linings create, newspaper and hotdogs never rot.
- (3) Because many substances, including newspapers and hotdogs, break down very slowly when covered with landfill linings.

Q3: What makes modern landfills less dangerous to the environment than older landfills? Select the best answer by circling its number.

- (1) The covering and linings make modern landfills less dangerous to water supplies than older landfills.
- (2) There are fewer bacteria and fungi to make materials decompose in modern landfills than in older landfills.
- (3) Recent excavations have found that materials do not decompose as rapidly in older landfills as they do in modern landfills.

Name: _____

Directions: What is the outcome of Kennedy’s mentioning the risk of nuclear war? Try to give an explanation to the above question while reading the passage below. After reading only once, write your explanation in the blanks below the passage.

The Cuban missile crisis of 1962 was the most terrifying confrontation of the Cold War. American aerial photographs taken revealed that the USSR had begun to place what US President John F. Kennedy considered offensive missiles (missiles designed to attack another country) on Cuban soil. Kennedy went on nationwide TV to tell the American people about the missiles and to demand their removal. He declared that the United States would not shrink from the risk of nuclear war and announced a naval “quarantine” around Cuba to prevent Soviet Union ships from bringing in additional missiles.

Explanation: _____

Q1: What did Kennedy mean by “the United States would not shrink from the risk of nuclear war”? Select the best answer by circling its number.

- (1) Kennedy did not want to risk the lives of people in the United States over Cuba.
- (2) Kennedy would negotiate to prevent a nuclear war in any way he could.
- (3) Kennedy would risk a nuclear war for the removal of Cuban missiles.

Q2: Why did the author refer to the Cuban missile crisis as “the most terrifying confrontation of the Cold War?” Select the best answer by circling its number.

- (1) Because it would increase the likelihood of nuclear war affecting the whole world.
- (2) Because it would increase the number of nuclear weapons set up by both United States and the Soviet Union.
- (3) Because it would decrease the chance of removing Cuban missiles.

Q3: What did the US speculate the Soviet Union was planning to do? Select the best answer by circling its number.

- (1) The Soviet Union was planning to point their nuclear missiles at Cuba.
- (2) The Soviet Union was planning to attack the US from Cuba.
- (3) The Soviet Union was planning to reduce the number of their nuclear weapons in Europe by bringing them to Cuba by ships.