UNDERSTANDING THE ROLE OF THE EPISODIC BUFFER OF WORKING MEMORY IN INFERENTIAL READING COMPREHENSION IN L1 AND L2 READERS UNDER VARYING CONDITIONS OF COGNITIVE LOAD AND DOMAIN KNOWLEDGE

by

MANPREET KAUR RAI

BSc. University of Calgary, 2004
M.S. Fort Hays State University 2006

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2014

Abstract

In recent years, Baddeley (2010) has added a new component, the episodic buffer, to his Working Memory (WM) model. The episodic buffer binds information from long-term memory (LTM) to the central executive but has been researched very little, especially with respect to its use with a second language. In fact, Juffs and Harrington (2011) stated, "To date there has been no research on the possible role of the episodic buffer in L2 learning and use" (p. 140). One goal of this study was to do just that.

Domain knowledge (DK) in baseball (Experiment 1) and English proficiency levels (Experiment 2) were used as proxies for difficulty level to study how inference processing under different conditions of domain knowledge and cognitive load in native (L1) and non-native (L2) English readers contribute to understanding the episodic buffer. In Experiment 1, 67 participants varying in domain knowledge about baseball read stories related to baseball with or without a concurrent cognitive load task of responding to tones while reading; they then answered comprehension questions of varying degrees of inferential difficulty. In Experiment 2, three groups varying in English reading proficiency, split into groups based on their lexical decision task scores (72 native, 40 intermediate, 40 beginner readers) read general stories with or without cognitive load and answered comprehension questions requiring varying degrees of inferential difficulty.

Accuracy and Reaction Time (RT) were differentially affected by working memory (OSpan), cognitive load, and inferential complexity. In Experiment 1, greater DK explained variance in effectiveness (accuracy) and efficiency (RT) as inferential complexity increased. In Experiment 2 OSpan was needed even at lower levels of inferential complexity for beginning readers. Surprisingly, for both experiments, participants responded faster under cognitive load conditions, although not at the expense of accuracy. This suggests that the episodic buffer is important for different levels of DK and proficiency, especially as the task becomes more difficult. Theoretical and practical implications of these findings are discussed.

UNDERSTANDING THE ROLE OF THE EPISODIC BUFFER OF WORKING MEMORY IN INFERENTIAL READING COMPREHENSION IN L1 AND L2 READERS UNDER VARYING CONDITIONS OF COGNITIVE LOAD AND DOMAIN KNOWLEDGE

by

MANPREET KAUR RAI

BSc. University of Calgary, 2004
M.S. Fort Hays State University 2006

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2014

Approved by:

Major Professor
Dr. Richard Jackson Harris

Abstract

In recent years, Baddeley (2010) has added a new component, the episodic buffer, to his Working Memory (WM) model. The episodic buffer binds information from long-term memory (LTM) to the central executive but has been researched very little, especially with respect to its use with a second language. In fact, Juffs and Harrington (2011) stated, "To date there has been no research on the possible role of the episodic buffer in L2 learning and use" (p. 140). One goal of this study was to do just that.

Domain knowledge (DK) in baseball (Experiment 1) and English proficiency levels (Experiment 2) were used as proxies for difficulty level to study how inference processing under different conditions of domain knowledge and cognitive load in native (L1) and non-native (L2) English readers contribute to understanding the episodic buffer. In Experiment 1, 67 participants varying in domain knowledge about baseball read stories related to baseball with or without a concurrent cognitive load task of responding to tones while reading; they then answered comprehension questions of varying degrees of inferential difficulty. In Experiment 2, three groups varying in English reading proficiency, split into groups based on their lexical decision task scores (72 native, 40 intermediate, 40 beginner readers) read general stories with or without cognitive load and answered comprehension questions requiring varying degrees of inferential difficulty.

Accuracy and Reaction Time (RT) were differentially affected by working memory (OSpan), cognitive load, and inferential complexity. In Experiment 1, greater DK explained variance in effectiveness (accuracy) and efficiency (RT) as inferential complexity increased. In Experiment 2 OSpan was needed even at lower levels of inferential complexity for beginning readers. Surprisingly, for both experiments, participants responded faster under cognitive load conditions, although not at the expense of accuracy. This suggests that the episodic buffer is important for different levels of DK and proficiency, especially as the task becomes more difficult. Theoretical and practical implications of these findings are discussed.

Table of Contents

List of Figures

List of Tables

Acknowledgements

"Now all glory to God, who is able through this mighty power at work within us to accomplish infinitely more than we might ask or think." (Eph 3:20). This dissertation is a testimony to the awesome power of God who has done infinitely more than I could ask or think in completing this journey to a PhD**.** I thank many individuals that have provided me with great assistance and guidance during the whole dissertation process. Without you, this dissertation would not have been possible. First and foremost I want to thank my advisor Dr. Richard Harris. Through all the ups and downs you have been there to help me grow in my academic development. Thank you for all the opportunities you have given me to succeed. Along with Dr. Harris, I thank the rest of my dissertation committee: Dr. Lester Loschky, Dr. Heather Bailey, Dr. Michael Harrington, and Dr. Lotta Larson. Your continued enthusiasm for the potential of my work is humbling.

The entire Department of Psychological Sciences, Kansas State University, has provided me with great support during my time here as a graduate student. I appreciate the entire faculty for doing all you can to help students succeed—special thanks to Department head, Dr. Michael Young for countless times of 'dropping in' for statistics help, Dr. Don Saucier, Dr. Jim Shanteau, Dr. Gary Brase, Dr. Mary Cain, and Dr. Laura Brannon for always being available for even the smallest of things. Thank you to Ila Axton, Adam Adair and Lacey Loken for keeping me on the right path as you answered all my silly administrative questions. Thanks Cathie for always making sure that I didn't get lost under all the recycling paper.

A word of thanks to my friends, peers and fellow colleagues in the department, Heather Hilgenkamp, April Park, Margo Woller, and Russ Webster to name a few. You kept me going in

Bluemont Hall at all hours of the day and night. A special thanks to my fellow graduate students in the Harris lab, Bernie de la Garza, Patricia Barros, Jorge Piocuda, John Smyers and Elena Knyshev and to the undergraduate research assistants, Seth Drees and Lindsay Cook. You helped sort out my dissertation when it was barely an idea to its completed version.

I wish to express my sincerest gratitude to Dr. Carol Shanklin, Dean of the Graduate School. You are not only the greatest Graduate Student Council (GSC) advisor, but also a wonderful person who truly cares for each graduate student at K-State. My professional development has flourished because of your support. Thanks to the GSC peers Kellan Kershner, Bethany Murray, and Taylor Wadian for your impact in my time here at K-State.

My heart beats for the Teaching and Learning Center, my GTA workplace, which was my place of refuge and solace during my dissertation writing. Thank you Dr. Jana Fallin and Dr. Vicki Clegg. It is because of you that I know what I want to do 'when I grow up.' Your love, prayers and support have made me into a better woman and professional. Thanks to Paula Connors, Nova Nickerson, Shay Dodson, Whitney Jeter, Cassie Wolfe, and David Fallin for making every day of my work something to look forward to. I appreciate you all than words can say.

I especially want to thank the English Language Program (ELP), specifically Abby Franchetti, and Mary Wood for allowing me to work with the ELP students. Thanks to all the ELP instructors and students that were patient with me and willing to try something new while helping me complete a dissertation. Without you, I would not have a study at all.

I owe my gratitude also to my fab five circle of friends: Vuyiswa Bushula-Njah, Junli Zhang, Arleta Ndlela, Mercy Palamuleni, and Caitlin Kelley who made me laugh when I wanted to cry, as we all walked the journey of grad school. A special thanks to Vuyiswa 'my person.'

You reminded me to focus on God instead of my circumstances, you stood with me when others wouldn't. Thanks for being my number one cheerleader and coach at same time, even if it meant doing one more all-nighter to get a section of the dissertation done. I am blessed and proud to call you my sister. I also thank little Iwhumni Njah, my 'third niece' for all your love, giggles, and wishing auntie Minnie sweet dreams.

I thank the many other Kansans who provided support from all ends of the state, from Beth Bills and the family in Hays, KS, to the Grace Baptist Church family in Manhattan. Your kindness and hospitality will remain with me wherever I go.

My deepest gratitude is reserved for my family. Thank you to my mom Balbir Rai. I will spend a lifetime in reverence and thanks for all of the sacrifices you have made for me and the entire Rai family. May you always know that your blood, sweat, tears and distance from your daughter was not in vein. My life and education is because of you. Thanks to my late-dad Gursharn Rai, who also sacrificed all for me. Thanks to Hardeep Rai, the best big brother a girl could ask for. I may be his dorky sister, but am blessed that he always has my back, even if I don't always make the best decisions. Thanks for being loving and patience even when we were separated by miles. Thanks to my sister-in-law (bhabi ji) Ranjeet Rai, my nieces Cassandra and Jasmine Rai, and nephew Deven Rai for reminding me that there is always something to smile and laugh about. You bring joy to my life everyday. I also want to thank by name the Shergill family (my aunt Kamal, uncle, Kundan, and cousins Mandeep and Satwant. You are all so special and important to me. I thank you for thinking of me and calling me exactly when I needed to hear an encouraging word. I will love you always.

Finally, I want to thank my future colleagues in D'Youville College in Buffalo NY, Dr. Todd Walter, Dr. Christine Walawander, and Dr. Patricia Abbott. Nothing helps a student finish

a PhD than a job waiting for them on the other side. Thank you for your encouragement and support in the last stretch of my degree. I know my time there will be just as blessed as it has been at K-State. Yes, I officially bleed purple.

Dedication

I dedicate this dissertation to my late dad, Gursharn Singh Rai. Your voice keeps playing in my head saying “Minnie doctor or doctor Rai”… I can feel your love and support even now. You left us too soon daddy, but think that you are smiling down from heaven knowing that your baby girl finally finished. You always knew I would. I love you and miss you every day.

Chapter 1 - Introduction

One may think that is silly to compare baseball domain knowledge to English proficiency, but I would argue this 'just ain't so.' Most would argue they are qualitatively different from each other, but as readers will understand, both require the retrieval of information from Long-Term Memory, which may be then processed via the episodic buffer of Working Memory (WM).

In recent years there has been much research conducted on Working Memory (WM) as a theoretical construct to understand higher level cognitive processes (Baddeley & Hitch, 1974; Baddeley, 1986; Baddeley 2000; Baddeley 2003; Baddeley, 2012). In addition, much work has also been conducted on WM as individual differences variables (such as WM span/capacity) in its relation to higher-level processes such as inference generation and the comprehension of those inferences (Calvo, 2001; Darke, 1998b; Linderholm, & van den Broek, 2002; Rai et al., 2011; Rai et al., in press; St. George, Mannes & Hoffman, 1997) in the context of reading extended discourse in both one's first language (L1) as well as their second language (L2) (Hammadou, 1991; Horiba, 1996; Kembo, 2001; Shimizu, 2005). There are several factors that have also been found to be important in the manifestation of these relationships between WM and inference processing, such as the effects of stress (Beilock & Carr, 2005; Eysenck et al., 2007; Ikeda, Iwanaga & Seiwa, 1996; MacIntyre & Gardner, 1994a, b).

The following pages will be an attempt to expound upon some of the theoretical explanations of the relationships mentioned above, as well as to discuss some other theoretical explanations of these relationships that may exist. Furthermore, important methodological approaches used to understand these relationships will be discussed in an attempt to suggest research possibilities that can be conducted with the specific variables mentioned above. This will help to fully understand the more recent additions to this WM construct, namely the episodic buffer, which has several grey areas with respect to its role in WM functioning, whether it be automatic or requiring cognitive effort, as well as its relationship to language processing. Since the episodic buffer's conception in 2000, there have only been a handful of studies that have been conducted to better understand its functioning, its existence, and the role it plays in overall working memory processes. This study is an attempt to elucidate some of the issues surrounding the episodic buffer, namely to see how the functioning of the episodic buffer can be used to better understand inferential comprehension in reading in L1 and L2 readers of English. Based on the literature, complex reading tasks are certainly those which should require the episodic buffer, although there have not been any empirical studies to assess this, especially in second language reading. In other words, as Juffs and Harrington (2011) state, "To date there has been no research on the possible role of the episodic buffer in L2 learning and use" (p. 140). One of the goals of this paper is to do just that, to understand the role of the episodic buffer in L2 learning, specifically in that of inferential processing while reading.

Baddeley's Working Memory Model

An understanding of the overall WM model is important to understand how the episodic buffer works. In general Baddeley and Hitch's (1974) (revised in 2000) model of WM is one in which information is temporarily processed, manipulated and stored until it is either forgotten or transferred to long-term memory. In other words, WM is a system which can both store and manipulate information (Baddeley, 2003; Baddeley, 2007). The key to understanding WM is that it is restricted by capacity and time. A person can remember a certain amount of information for only a brief time, approximately one minute, unless that person puts some sort of cognitive effort into remembering that item (Baddeley, 1986; Baddeley, 2000). As such, this limited time constraint suggests that the faster that information flows through components of WM, the more that can temporarily be stored in WM. For instance, proficient readers of English can read more quickly, and therefore, the flow of words that they must keep in mind does not strain their WM as much as a non-proficient reader of English, whose flow of information may not be smooth, and therefore the L2 readers are unable to maintain information efficiently (Ardila, 2003; Daneman & Carpenter, 1980; Harrington & Sawyer, 1992; Juffs, 2004; Leeser, 2007; Miyake & Friedman, 1998; Walter, 2004). This notion of the amount of information that can be actively processed and stored simultaneously in a limited amount of time is known as WM span (WM capacity) (Daneman & Carpenter, 1980). In this sense, WM is an individual differences variable, as capacity can vary (Daneman & Carpenter, 1980; Unsworth et al., 2005). Span can be measured in terms of accuracy (effectiveness of comprehension) or the speed at which information is processed (efficiency) (Eysenck et al., 2007). WM itself is comprised of four major components. There is the central executive and three "slave" systems working under the central executive. These slave systems are known as the phonological loop and the visuospatial sketchpad (VSSP), and a new component known as the episodic buffer.

The central executive (CE) is the important component that acts as the intersection through which all information must pass in order to be processed, regardless if it is phonological or visual in nature. Thus, the CE processes is an attention-based conscious system. With respect to language processing, the phonological loop (which is the temporary store for verbal information), episodic buffer and the central executive are of most concern (although the VSSP is needed to actually read the words). In fact, Miyake et al., (2000) postulated that the phonological loop is important in learning new words by adults, as well as children. Others have also agreed with this in stating that the phonological loop is crucial for language acquisition in learning any language (Perani, 2005; Ardila, 2003). WM itself, particularly the function of the phonological loop, is also correlated with many language properties such as "the ability to comprehend written and spoken words, interpret the intended meaning of an ambiguous word, use context to infer the meaning of an unfamiliar word, and produce category exemplars, synonyms, and sentences from given words" (Michael & Gollan, 2005, p. 399). Once these individual words are processed, however, it is the CE that must actively maintain the individual words online, as well as integrate the meaning of the words together into surface level meaning, and as any inferential meaning that the text may have. Thus, although the phonological loop is of great importance, it is the CE that drives its functioning. However, it has been found that in many cases, individuals are able to store much more information in WM, especially with respect to long prose, which is well beyond the capacity limits of the phonological loop, in terms of 7 +/- 2 items (Baddeley, 2000; Baddeley & Andrade, 2000; Ericsson & Kintsch, 1995). As such, in order to provide a coherent explanation of the role of WM in comprehending extended discourse, Baddeley (2000; 2003; 2007) has proposed a 3rd slave system (or 4th subcomponent of WM), referred to as the episodic buffer.

This buffer is like the other two slave systems in having a limited capacity (approximately 4 chunks of code) storage system; however, beyond storage it also has some processing abilities, although those abilities are controlled by the central executive (CE) (Baddeley, 2000; Baddeley, 2007). The items that can be represented in the episodic buffer can be multidimensional, integrating both phonological, as well as visual, information. Furthermore, it is linked to long term memory (LTM) which provides it input to be passed on to the CE. Therefore, it has links to both of the other two transient storage systems (phonological loop, visuospatial sketchpad), a link to the transient CE, as well as a link to the stable LTM (such as episodic LTM, one of the many types of LTM, all of which can be linked to the episodic buffer)

(Baddeley, 2007). As such, the episodic buffer is "episodic in the sense that it holds episodes whereby information is integrated across space and potentially extended across time" (Baddeley, p. 421), so it can draw from resources in episodic LTM. It is a buffer in that it acts as a buffer between the other two slave systems, as information in those two systems coded in different dimensions can interact. This interaction and active linking to episodic LTM leads to a key component of the episodic buffer, namely that it operates through conscious awareness in order to bind (integrate) information from different sources into coherent chunks (Baddeley, 2007). Furthermore, this binding of information can be passive (static) where two repeatedly co-occurring items (i.e. 'yellowness' + 'bananas') can be bound with little attentional effort. Binding can also be active (dynamic) in which novel items that do not co-occur can be combined in a number of different combinations (i.e. a purple banana stuck in snow cone pot). This type of binding does require more attention (Baddeley, 2003). The episodic buffer can do both, although again it is limited by how much attentional control the CE is exerting at the same time.

As such, the episodic buffer can essentially bind information from the subsystems of WM as well as long-term memory into a unitary episodic representation (Baddeley, 2000, 2003). Specifically, in the process of reading long prose, the central executive uses the goal-directed attentional system to retrieve knowledge from long-term memory to the episodic buffer and process it there for meaning. If WM resources become strained (such as in stressful situations), the stimulus-driven (bottom-up) attentional system gains precedence because the central executive is unable to make full use of its resource-intensive goal-directed processes (Eysenck et al., 2007; Gray, et al., 2002); (Rai et al., 2011, p. 190). Specifically, because the episodic buffer can simultaneously retrieve semantic representations from episodic LTM, while holding several chunks (episodes) from integrating prose, a greater number of coherent sentences can be held in storage, thereby enabling the CE to manipulate those sentences at a later time in order to comprehend them. Thus, the ability for extensive binding to occur allows the comprehension of extensive prose. That being said, if the CE is unable to communicate with the episodic buffer, then a person may not be able to comprehend efficiently. The episodic buffer will be discussed further shortly.

Working Memory and Reading Comprehension

WM is a crucial component in language comprehension (Alptekin & Ercetin, 2010; Ardila, 2003; Daneman & Carpenter, 1980; Harrington & Sawyer, 1992; Juffs, 2004; Juffs & Harrington, 2011; Leeser, 2007; Miyake & Friedman, 1998; Rai et al., 2011; Walter, 2004). Specifically, during the comprehension of extended text, the two components of WM that play a major role are the central executive (CE) and the episodic buffer (Baddeley, 2000; Daneman & Carpenter, 1980). For instance, when reading complex text, especially those that require the drawing of complex inferences (such as knowledge-based pragmatic inferences), these inferences increase the demands on the CE of working memory, thus, impairing reading comprehension (Calvo, 2001; Linderholm & van den Broek, 2002). Further details of inference comprehension during reading will be discussed shortly. Specifically, the CE is the ‘mainframe’ of WM as it actively processes and manipulates information held in the slave systems, regardless of the modality (visual, phonological, or combination of both) of the information. The episodic buffer serving as a ’slave’ system to the CE integrates information that is stored in various modalities (visual, phonological or semantic) so that it can be later processed by the CE. The buffer is also able to act as a relay station, or buffer, between long-term memory (LTM) and the CE (Baddeley, 2000; Baddeley, 2003).

As such, when reading extended text, WM uses both bottom-up (stimulus-driven) and top-town (goal-directed attention) systems to extract meaning from the text. Specifically, the goal-directed attentional system is used to retrieve knowledge from long-term memory via the episodic buffer, after which the CE can process it. If WM resources are strained, the CE can not make use of the goal-directed process (which require considerable resources). When this happens, the stimulus-driven attentional system comes into play. (Eysenck et al., 2007; Gray, et al., 2002; Rai et al., 2011). In other words, when reading, the CE not only maintains what has been read online, but also manipulates that information in a meaningful way by integrating the surface-level text with higher discourse processes (such as inferential processing) using the episodic buffer for extended discourse. However, if the reader has comprehension difficulties by their inability to process those higher situation model discourse processes, then the CE reverts back to relying more on the stimulus-driven text base levels to comprehend the text.

So, if a reader is viewing a text about a restaurant scene where a waiter goes to get a broom and dustpan after a pitcher falls on the ground, they must 1) hold the propositions of the textbase level in WM, while 2) retrieving relevant world knowledge from long-term memory into WM (e.g., pitchers are usually made of breakable material such as glass, which shatters when it falls on the floor), and 3) use that retrieved knowledge to fill in the missing propositions needed for the situation model to make sense (e.g., the fallen pitcher must have shattered, causing the waiter to get the broom and dustpan *in order to clean up the glass*) (Rai et al., 2011, p. 189). In fact, when creating the reading span task to assess reading comprehension and WM span Daneman and Carpenter (1980) found a positive correlation between WM span (processing and storage capacity) and reading ability. This correlation between WM and reading was found at all levels of inferential processing from answering the least difficult non-inference (fact) questions, to those of greater difficulty: pronoun reference questions (involving bridging inferences), and topic summary questions (involving pragmatic inferences) after reading short passages. Taken together, this suggests that WM indeed plays a crucial role in reading comprehension in general, and especially inferential comprehension to which we will explore further next.

Working Memory and Inference Comprehension

Only a handful of studies have investigated the role of WM in inference comprehension, although there has been disagreement as to the taxonomies of inferences used. The varying taxonomies (Graesser, Singer & Trabasso, 1994) of inferences have one commonality that they go from least complex (i.e. facts that would be non-inferences consisting of items directly stated in the text), to more complex (i.e. bridging inferences such as pronominal and anaphoric referents that are in the text, but must be held in memory and referred back to), to most complex (such as pragmatic inferences that are not mentioned in the text, so the reader must use their long-term memory stores as well as WM functioning to integrate the text with unstated information) (Graesser, et al., 1994; Kispal, 2008; Singer, Harkness, & Stewart, 1997). Specifically, readers may use several levels of mental representations of the text in order to comprehend inferences. Kintsch (1994, 1998) discusses four levels of representation of knowledge from text in memory. The shallowest (surface) level involves encoding the exact words, each with its particular string of phonemes. The second (textbase) level involves processing on a sentence or utterance level. This involves the specific propositions constructed by the hearer or reader during comprehension. The third (situation model) level draws on the reader or listener's extensive knowledge base about topics mentioned in the text or utterance. Finally, the pragmatic level involves social and contextual constraints.

According to the constructivist approach of inference processing (Graesser, et al., 1994), readers are continuously engaged in a constant 'search (effort) after meaning' to build a situation model of the text that is coherent both at the local (textbased) and global (situation model based) level. In this way they can draw all the inferences needed to explain why things are mentioned in the text in order to achieve coherence (Graesser et al., 1994). More specifically, it is argued that of a possible 13 classes of inferences, those that establish local coherence are processed online, and those that elaborate on the text beyond the explicit text alone are processed offline (after reading). Thus, the search after meaning can occur across several points during and after reading. However, this search after meaning can also break down under situations in which the individual does not make any effort to establish coherence, if they lack relevant background knowledge to deal with breaks in coherence or are not given sufficient time for comprehension (Graesser et al., 1994; Marmolejo-Ramos et al., 2009). This is consistent with other studies that have shown that inferences are not drawn under conditions of time constraints or when an extra load is placed on WM (Eysenck et al., 2007; Rai et al., 2011). This constructivist model is also consistent with the resonance model (Myers & O'Brien, 1998), which postulates that "generation of inferences begins with automatic activation of elements in long-term memory that are similar to information being comprehended. If there is a sufficient match, the controlled phase is initiated, and items in long-term memory are pulled into working memory where they are integrated into a mental representation" (Blake, 2009a, p. 361). This idea of integrations of mental representations fits well with the function of the episodic buffer.

Thus inferences drawn at the levels requiring larger units of discourse to be processed (situation model) are generally harder than those requiring lower levels (such as the textbase level) (Kinstch, 1998). Furthermore, inferences drawn at the textbase level usually involve connecting two propositions together (i.e. in "Bob went to the store where he bought apples" one must connect 'he' with 'Bob'). This inference is referred to as a bridging inference (or pronominal reference). At the higher level processes, such as situation model, inferences are made using the reader's background knowledge to fill in missing information not explicitly mentioned in the text. For instance, "Bob went to the store and was soaking wet by the time he arrived." Here, one must use their background knowledge of reasons why someone would be wet to infer why Bob got wet on his way to the store, perhaps because it was raining. This inference in referred to as a pragmatic inference (Harris & Monaco, 1978; Johnson-Laird, 1993; Kispal, 2008).

Several studies have shown that readers with lower WM capacity are less likely to draw pragmatic inferences (Calvo, 2001; Linderholm & van den Broek, 2002; Rai et al., 2011; St. George, Mannes & Hoffman, 1997), or require more processing time to do so (Estevez & Calvo, 2000; Rai et al., 2011) than readers with higher WM capacity. Also, readers with lower WM capacity are less effective in identifying the referent of a pronoun (a bridging inference) or identifying the general theme of a story (a global/pragmatic inference) (Daneman & Carpenter, 1980). Together, these results suggest that WM capacity may become more important as the degree of inferential processing required for reading increases, which converges with the constructivist approach suggesting that more effort is needed for more difficult inferences; thus having a greater WM capacity allows greater resources to be allocated to those more difficult inferences.

Working Memory and Inference Comprehension: L2 Reading

Several studies have also found there to be positive correlation between L2 proficiency and the likelihood of drawing higher level inferences during reading, with those who have higher L2 proficiency drawing more inferences and those with lower L2 proficiency being less able to draw inferences, or only lower level (bridging) ones (Horiba, 1996; Kembo, 2001; Shimizu, 2005). Furthermore, studies have found that when those with lower L2 proficiency draw inferences, they do so with less accuracy (and sometimes more time taken to respond) than those with higher L2 proficiency (Rai et al., 2011). As noted earlier, the level of inferential processing during reading can be constrained by executive WM resources (Calvo, 2001; Daneman & Carpenter, 1980; Estevez & Calvo, 2000; Linderholm & van den Broek, 2002; St. George et al., 1997), especially for those with lower foreign-language (FL) proficiency (Harrington & Sawyer, 1992; Service et al.,2002; Van den Noort et al., 2006), for whom resources are more taxed than for those with higher FL proficiency. This is further illustrated by Hammadou (1991) who found that those with lower FL proficiency typically make more incorrect inferences (i.e., unsupported by the text) than those with higher FL proficiency. She also found that having familiar text (based on topics readers are familiar with) aids those with lower L2 proficiency to make inferences suggesting that even those with lower L2 proficiency are able to draw the (pragmatic) inferences that use the situation model of representation (Horiba, 1996; Kintsch, 1998; Shimizu, 2005) under certain situations. These data suggest that as L2 proficiency increases, becoming more automatized, it strains the CE less, allowing it to have more resources to process more complex inferences (Rai et al., 2011). This is not a surprising claim, since these studies do show that WM and foreign language proficiency are correlated, with an increase of both resulting in processing at higher levels of discourse processing, such as those of pragmatic inferences (Rai et al., 2011).

To further clarify this point, Blake (2009b) found that those with right hemisphere brain injury had difficulty with comprehending discourse, as they were unable to use contextual cues while reading in certain conditions (similar to those with low WM span, or low L2 proficiency). However, when they were given extra time, along with the repetition of highly predictive outcomes of a text, then they were able to generate the appropriate predictive inferences (Blake, 2009b). Highly predictive outcomes may be a sentence that claims a person 'will' perform an action whereas low predictive items may be a sentence that claims a person 'may' perform an action. Blake's (2009b) study also found that those with right hemisphere damage and low WM spans had even further deficits in their ability to make inferences, even when outcomes were quite predictable. Thus, this evidence again converges with other findings, suggesting that capacity limits can decrease cognitive processing abilities in the CE, in both one's L1 and L2, which will now be discussed further.

Working Memory and L2 Comprehension

Working Memory is important for reading in a first language, but there is also a correlation between reading proficiency in one's first language (L1) and reading proficiency in one's second language (L2). Keeping this in mind, Osaka and Osaka (1992) conducted a reading comprehension study using Japanese-English bilinguals who had studied English for six years and were relatively proficient in the language. Using the reading span test in both English and Japanese, the researchers found that second-language processing may draw from the same WM resources as L1 processing, since L1 reading span scores were highly correlated with Japanese reading span scores. These results were replicated by Osaka et al. (1993) using a sample of French-English bilinguals. A similar study comparing L1 to L2 to L3 proficiency levels to WM functioning in three languages (Van den Noort et al., 2006) found correlations between proficiency and reading spans in the three respective languages suggesting that "a constant underlying WM capacity regardless of the language one is using" (Rai et al., 2011, p. 192), even as WM resources decrease if proficiency in an L2 and L3 is low.

In another study using Japanese-English students writing the Test of English as a Foreign Language test (TOEFL), Harrington and Sawyer (1992) found that L2 readers with high WM spans scored higher in both the reading and grammar sections of the TOEFL when they had had high WM than those with low WM spans. Similarly, Juffs (2004) conducted a comparative study of Chinese and Japanese bilinguals, using English native speakers as a comparison group. All three groups were given ambiguous and unambiguous English sentences. Of the three groups, the Japanese were the least proficient in English, and as such they also scored the lowest on the WM span tests as well. The low WM span individuals also had the slowest reading times. Thus, proficiency and WM spans seem to go hand in hand. In another study assessing reading comprehension, Hulstijn and Bossers (1992) found that when first learning to read an L2, factors such as vocabulary and grammar rules guide reading. As readers become more proficient, in later learning WM and L1 proficiency begin to play a role in reading as well; these are known as L2 non-specific factors. This suggests then that individual differences may arise in reading comprehension, but more importantly that WM does play a role in language proficiency.

In another important study putting L1 and L2 proficiency and WM together, Miyake and Friedman (1998) proposed that linguistic cues in L2 act as mediators of the effect of WM on L2 syntactic comprehension. This is based on the fact that different cultures use (prefer) different cues when comprehending sentences that they are reading. For example, English speakers use more word order cues, whereas morphological agreement is more important for Italian, French, and Spanish speakers, and case markings are important for Japanese and Germans, who also frequently use animacy (Miyake & Friedman, 1998). Thus in a phrase such as "the pencil the cows are kicking" English speakers would initially pick the pencil as the agent whereas Japanese speakers would choose the cow as the agent (Miyake & Friedman, 1998 p. 349). It is assumed that WM plays a role in this acquisition of cue preferences. Thus, in order to test how WM influences syntactic comprehension and how cue preferences mediate this relationship, Miyake and Friedman (1998) gave Japanese-English speakers a WM reading span test in both L1 and L2, a cue preference test, as well as a syntactic comprehension test in English, in which the participants indicate who did what to whom in a sentence. This syntactic comprehension is an example of a low-level text-based (bridging) inference. The researchers found a correlation between performance on all four measures. A path analysis further found that L1 WM determines L2 performance, which makes sense as the same resources are being used in both WM tasks. Also, L2 WM determines cue preference distance, indicating that those with higher WM spans are better at comprehending complex sentences. The third path in the path analysis indicated that L2, WM, and cue preference distance determine syntactic comprehension, which makes sense also, because Japanese participants had to switch to using word order cues for English sentences, so those who had trouble making the switch did not understand the sentences in the syntactic comprehension task . The final path showed that L1 has an indirect effect on both cue preference and syntactic comprehension. This alludes again to the individual differences, in that cue cost that may be taking place. That is, cues can place demands on effective WM processing. This cue cost may come about because word order cues place a higher demand on WM processing than do morphological cues, and as such Japanese participants have a more difficult time in syntactic comprehension in English than English speakers may have reading Japanese sentences (Miyake & Friedman, 1998). This further shows that high WM span Japanese bilinguals are closest to native English speakers in their cue preferences, as they are influenced by both animacy and word order information, whereas low WM span bilinguals are influenced mostly by animacy. Again with respect to language learners, those who are at lower levels of English proficiency may have a larger problem when comprehending text because written Japanese is a syllable-based language, in contrast to the alphabetic form of English; thus processing it may take an even larger toll on WM. Taken together, all of these studies demonstrate that WM plays roles in various aspects of foreign language learning, especially with respect to reading.

WM and Cognitive Demands

As we have seen, cognitive demands can put a strain on the CE of working memory resulting in poor performance on cognitive tasks. One such theory attempting to account for this is the distraction theory (Wine, 1971) that postulates, that under high cognitive load situations (i.e., stress/pressure), individuals who would normally perform well under no stress conditions end up performing quite poorly (Baumeister, 1984). The reason that the distraction theory can describe this decrease in performance is that cognitive tasks such as mathematics, or inference generation in comprehension, place a high cognitive demand on WM, especially the central executive needed for attention-based executive functions. When more demand is introduced to the already demanding task, the individual may have anxious thoughts which "distract a person's attention from the task, thus depleting the limited executive resources needed to complete it (Ashcraft & Kirk, 2001; Baumeister, 1984; Beilock & Carr, 2005; Rai et al., 2011).

As an extension of the distraction theory, Eysenck et al, (2007) have proposed the Attentional Control Theory (ACT). This theory is based on the premise that "the effects of anxiety on attentional processes are of fundamental importance to an understanding of how anxiety affects performance" (Eysenck et al., 2007, p. 338). More specifically, the ACT puts forth the notion that anxious thoughts distract attention away from one's current goals (such as reading for comprehension, or speaking in public). Thus, instead of goal-driven attention (which enables the inhibiting of irrelevant information, shifting attention, and updating information in working memory), anxiety increases the influence of stimulus-driven attentional control, directing attention to threatening stimuli or thoughts and reduces the influence of one's attentional control. As aforementioned, the CE is responsible for active processing and manipulation of information. As such, according to the ACT, for those who have lower WM capacity, the distracting effects of anxiety are greater than for those with higher WM capacity. Specifically, in performing stressful (or cognitively demanding) tasks, there seems to be a tradeoff of efficiency (speed of performance) and effectiveness (accuracy), when the central executive is under such taxing conditions (e.g., anxiety) that the attentional resources needed to perform a task are strained, thus impeding performance (Eysenck et al., 2007; Rai et al., 2011). This difference in performance under stress is further increased when the task is complex, as the extra stress further strains an already difficult problem. However, to compensate for lowered processing abilities in effectiveness, individuals can increase processing time (i.e., decreasing efficiency). Some compensatory strategies may include regressive eye movements in the text, and sub-vocal rehearsal, both of which decrease efficiency. "According to this view, anxiety will more commonly lead to decreases in efficiency than in effectiveness, so long as the person has control over how quickly they respond," (Rai et al., 2011, p 194) also suggesting that if speeded responses are required, and the time to respond is fixed, compensation will not be able to occur. As aforementioned, this theory fits in with theories of inference comprehension (such as the constructionist model of Graesser et al. 1994), which suggest that as difficulty increases and if there is not enough time for an individual to adequately process text, then higher level (i.e. pragmatic) inferences will not be drawn.

Working Memory and Domain Knowledge

As aforementioned, fluent reading requires the access of information from long-term memory, especially when constructing the situation model. Furthermore, the episodic buffer binds information from long-term memory and the other slave systems as a temporary store of information to be later manipulated by the central executive, even when WM is highly strained. The assumption then that domain knowledge (DK) would contribute favorably to enable the episodic buffer to pull information from long-term memory (if there is sufficient background knowledge of the text in LTM) should not be surprising. However, this relationship between the episodic buffer, cognitive load, and domain knowledge when reading prose has yet to be examined empirically. However, much research has examined domain knowledge and working memory in general, especially with respect to the knowledge of baseball. A series of studies starting as early as the 1970's has established that, when reading stories about baseball, those with high DK comprehend the stories in more detail than those with low DK (Chiesi, Spilich, & Voss, 1979; Fincher-Keifer, et al., 1988; Hambrick & Engle, 2002; Ricks & Wiley, 2009; Spilich et al., 1979; Stahl et al, 1991). Generally, it has been established that domain knowledge facilitates reading comprehension and subsequent memory of that text (Daneman & Merikle, 1996; Recht & Leslie, 1988; Spilich, et al., 1979; Walker, 1987). Even studies of L2 language processing have suggested that knowledge of the topic can help readers perform better than those with little topic knowledge (Guo & Roehrig, 2011; Miller & Keenan, 2011). Miller and Keenan (2011) found that L2 readers with adequate domain knowledge of the text were able to recall as much central information and peripheral information as their L1 counterparts. However, recall deficits in L2 readers were relatively greater for central information than peripheral information in the story, due to resources being allocated to text base level reading rather than to the situation model. This is known as the centrality deficit. In fact, DK is treated as a skill that can be practiced so that as knowledge is being acquired (early stages of skill), more attention and cognitive resources are needed to acquire that knowledge, but once the knowledge is proceduralized, then that is less important as skill is acquired. This is known as the circumvention-of-limits hypothesis (Ackerman, 1988; Hambrick & Meinz, 2011). This, as cited in Ericsson and Charness (1994, p. 725), "Performers can acquire skills that circumvent basic limits on working memory." If we think about this in terms of language processing, those with greater knowledge of the text, or reading skills in general, should theoretically have the 'acquired knowledge' for the language (i.e. syntactic knowledge, grammar rules), and therefore are able to create situation models with much more ease than those with less knowledge of a language. Similarly, if more attention is needed in earlier stages of acquiring a language, it may be the case that the episodic buffer plays more of a role in performance in those earlier stages than in later stages.

Some studies have found that DK drives cognitive performance, regardless of levels of WM (Ackerman & Kyllonen,1991; Britton et al., 1998; Fits & Posner, 1967; McNamara, 2004). In other words, "high levels of domain knowledge can compensate for lower levels of working memory capacity" (Hambrick & Engle, 2002, p. 342). Other studies have suggested that it is WM that drives cognitive performance regardless of DK, whereby those with higher WM have more cognitive resources to facilitate their DK (Cantor & Engle, 1993; Just & Carpenter 1992) similar to the 'rich-get-richer hypothesis' (Hambrick & Engle, 2002). Finally, similar to the previous model, there are those (Haenggi & Perfetti, 1994; Hunt, 1978; Posner & McLeod 1982), who suggest that although DK is important in cognitive performance, “working memory capacity is a basic mechanism underlying proficiency in cognitive domains” (Hambrick & Engle, 2002, p. 342). In an important study by Hambrick and Engle (2002), structural equation modeling was used to assess the directional effects of baseball knowledge, general knowledge, WM capacity, and memory performance on questions related to descriptions of baseball games. This study was done to better understand the conflicting results of the relationship between WM and DK. As such, Hambrick and Engle tested all three models. They found the WM capacity had a direct contributory effect on memory performance, independent of domain-specific knowledge alone (which had a main effect of performance in itself). This was also consistent with hierarchical regressions performed that indicated that working memory capacity predicted performance on baseball texts above and beyond domain knowledge alone. This is not to say that domain knowledge is not important in itself (as it predicted more than half of the variance in performance), but the non- interaction of WM and domain knowledge (and other studies that look at WM and domain knowledge interactions (Hambrick & Oswald, 2005; Hambrick & Meinz, 2011) seem to suggest that WM is what drives performance. In other words, “Activation of domain knowledge by the familiar context [does] not reduce, much less eliminate, the effect of working memory capacity on performance.” (Hambrick & Meinz, 2011 p. 277), but instead it appears that WM facilitates performance when one has adequate DK, which is consistent with the rich-get-richer hypothesis.

A Closer Look at the Episodic Buffer: Episodic Buffer and Long Term Working Memory

As aforementioned, much of the integration of prose using representations from LTM occurs with the episodic buffer. Some (Baars & Franklin, 2003; Franklin & Ramamurthy, 2006;

Was, 2014) have suggested however, that this buffer is nothing more than giving Ericsson and Kintsch's (1995) construct of Long Term Working Memory (LTWM) another name. Surprisingly, Ericsson and Kintsch have formally said nothing about the episodic buffer, although the proponents of the episodic buffer have always made it clear that it is a separate construct than LTWM.

LTWM is, as the construct implies, a stable temporary storage system within LTM from which retrieval can occur based on skilled use and attention-based retrieval cues (called retrieval structures) located in short-term memory (STM) (Ericsson & Kintsch, 1995). Furthermore, it is assumed that, once information is stored in LTWM, it can remain active indefinitely, so if a skilled task is disrupted, it can easily be continued with activation of the appropriate retrieval structures. It is important to note that the notion of LTWM especially applies to skilled tasks and experts (such as mnemonists or chess experts), who can hold several chunks of activated information beyond seven plus or minus two, and retrieval of these large chunks can occur quite quickly (within 300 msec). Specifically, Ericsson and Kintsch (1995) claim that "LTWM simply argues that subjects can acquire domain-specific memory skills that allow them to acquire LTWM and thus extend their working memory for a particular activity" (p. 214). Thus, those who have mastered a particular skill can retrieve information from LTM as quickly as from STM, because of the increased activation in retrieval cues. This has been shown in several cognitive tasks, such as using baseball knowledge (Ricks & Wiley, 2009), text production (Kellogg, 2001), as well as tasks showing that LTWM can help overcome interruptions using expertise skill such as reading (Oulasvitra & Saarlilumoa, 2006), as once items are encoded into LTWM they are safeguarded against forgetting.

In the context of inference processing, text representations generated during comprehension, based on Kintsch's (1998) levels of representation, become retrieval structures in LTWM, thereby creating additional cues by which later comprehension can occur. The key is that relevant material being read in previous sentences needs to be accessible and active in order to be comprehended. Because these retrieval cues are stable and can be reactivated easily as they are stored in STM, the role of LTWM is suggested to enable readers to be able to comprehend text, even when disruptions occur in the midst of reading, (as the retrieval cues can readily activate the previously read material). Studies supported that conclusion (Ericsson & Kintsch, 1995; Oulasvirta & Saariluomna, 2006), in that individuals who were interrupted while reading text showed no decrements in performance (in neither accuracy nor speed) when given subsequent comprehension questions, if speeded responses were not required. In other words, skilled readers are better able to access LTM while reading prose using retrieval cues held in activated portions generating LTWM. The problem herein is that, in terms of L2 processing, low proficiency individuals are not skilled readers, and therefore, this model (as the authors note) is more suited towards skilled activities rather than toward those where a high level of skill may not exist.

With respect to the episodic buffer, LTWM seems to have several similarities, such as that of integrating information from LTM to use on a current task. However, the nature of the episodic buffer is very transient and temporary with respect to retrieval, as it is dependent on the central executive and can be accessed immediately, without the apparent need of expert skill (Baddeley, 2000, 2007, 2012). LTWM on the other hand emphasizes the stable development of retrieval structures already within LTM, "as in the case of mnemonists who are capable to demonstrating remarkable immediate recall only because of many hours of practice building up the necessary structures in LTM" (Baddeley, 2007, p. 145). Furthermore, activation in LTWM is seen as just the reactivation of existing knowledge, whereas the episodic buffer is "capable of creating new structures, which can themselves be manipulated and reflected upon (Baddeley, 2000, p. 420). Specifically, in the examples of the dynamic binding above, and the famous example of Baddeley (2003) describing an elephant playing hockey, new representations of these items that typically do not occur together are made within the episodic buffer to form a coherent episode of it so that later comprehension (e.g., of questions such as what position should the elephant be) can occur with greater effectiveness and efficiency. Although the present studies do not directly test LTWM, they suggest that LTWM may not be best suited to conceptualizing reading in an L2, since LTWM assumes that novel representations are not made, but only prior knowledge drives performance. It is important to note that perhaps the two may have some overlap, despite their differences, which will also be discussed later.

Additionally, another WM model by Cowan (1988, 2001, 2014) asserts that working memory is not necessarily a separate entity in and of itself, but is the activated focus of attention as information moves to and from LTM. To illustrate, working memory (or the focus of attention) would be considered the yolk of the egg of memory (with the white part memory in general). Although all of these theories have different underlying mechanisms of how exactly memory works, it is clear that the episodic buffer (like the other theories) helps address the crucial communication between LTM and WM. Also, it is important to note that although these alternative theories are not the crux of the current studies, they do illustrate that they can still help put the episodic buffer in context in that it is the addition to Baddeley's model to better help explain the interaction between WM and LTM.

Episodic Buffer as a Real Construct

Much of the work done thus far has been to establish the existence of the episodic buffer, using dual-task paradigms (Baddeley, 2007; Morey, 2009; Was & Woltz, 2007). Still, there have been several studies (behavioral, neurological, and some physiological) to show that it is indeed a real construct distinctly separate from the CE and the other two slave systems. In neuroimaging work (Baddeley & Wilson, 2002; Berlingeri et al., 2008; Rudner & Ronnberg, 2008) examining both normal patients and those with dementia, it was found that, in tasks that integrated and required binding of multidimensional information, the left anterior hippocampus-binding information along with the prefrontal cortex (PFC) was activated, as opposed to areas activated by the phonological loop or the visuospatial sketchpad alone. More importantly, areas known to show activation of CE functioning (PFC) were activated but in different locations from the integration points, suggesting that the episodic buffer is a real validated construct. More recent studies (Berlingeri et al, 2008; Luck et al., 2010) have further found that the medial temporal lobe structures are activated. These are needed for the maintenance and encoding of bound information, but not necessarily needed for retrieval, consistent with the role of the episodic buffer of being an "integrative buffer dedicated to the formation and short-term maintenance of bound information being entered or retrieved from LTM" (Luck et al., 2010, p. 945). For instance, earlier studies of the episodic buffer, maintained that all binding occurs via the central executive actively processing information from the different stores using conscious attention (Baddeley, 2000, 2007). Later however, it has been demonstrated that the episodic buffer does not always need conscious attention but instead can serve as a passive store, like the other two slave systems, wherein binding is automatic (Baddeley, Hitch, & Allen, 2009; Karlsen, Allen, Baddeley, & Hitch, 2010). These studies have been done with items with visual features such as colors and shapes (Allen, Hitch, & Baddeley, 2009; Karlesen et al, 2010; Vogel et al., 2001), as well as chunking in sentence recall (Baddeley, Hitch & Allen, 1999), to name a few. Such studies have also been conducted with various populations of participants such as those with Alzheimer's disease (Huntley & Howard, 2009), those with ADHD (Alderson et al., 2014), and children (Henry, 2010; Sluuzenski, Newcombe & Kovacs, 2006).

Henry (2010), found that children with intellectual disabilities were able to perform as well as their mentally matched pairs on a story-recall task, paired-associate task, and a category task, all of which required them to bind information from their LTM with other sources of information (such as information from the other two slave systems), and since the performance is comparable, it is likely that the episodic buffer is able to hold this bound information automatically. Baddeley et al., (2009) conducted four experiments on sentence and word list recall using varying types of concurrent tasks and found that the chunking processes involved in sentence recall did not especially rely on the central executive, suggesting that the episodic buffer does not actively bind information (such as that needed for sentence recall). In other words, "chunks enter the episodic buffer where they become available to attention, but the binding process whereby chunks are formed and enter the buffer occur outside attention" where they can later enter into the focus of attention (Baddeley et al., 2009, p. 453). This further suggests that binding can occur even when the central executive is taxed, but at the cost of performance decrements on the specific task.

Similarly, Baddeley, Allen and Vargha-Khadem's (2010) studies with amnesic patients with trauma to the hippocampus found that the amnesics could perform as well as controls with verbal and visual materials, suggesting that the hippocampus (which typically is activated with conscious awareness) may not be necessary for binding as suggested by initial reports of the episodic buffer. Interestingly, however, Jeffries, Ralph and Baddeley (2004) conducted a dual-task study with a visual choice reaction task on recall of stories, sentences, and lists of unrelated words (thus decreasing complexity across all three types). Participants listened to either words, sentences or stories while simultaneously completing a choice reaction time task which displays boxes with the participant having to press one of 4 keys as quickly as possible when a star appears in one of the boxes. At the sound of a tone, they must recall what they heard in order. It was found that the dual-task interfered with the processing of sentences but not stories, since in stories chunks can be made based on gist parts instead of trying to chunk (relate) unrelated sentences together. This has two implications, the first being that story segments are more able to be chunked and integrated into the episodic buffer while making semantic representations in LTM. As such, the episodic buffer was present in this task. Also, this suggests that "the episodic buffer function is not necessarily reliant on the central executive function unless loop capacity is exceeded" (Rudner & Ronnberg, 2008, p. 22). This idea of the episodic buffer being generally reliant on the central executive, but under 'static' or 'low attention/processing' tasks is seemingly able to function on its own has also been found in other studies exploring tasks such as the N-back task for recall, and verbal recall tasks (Baddeley, 2007; Baddeley, Hitch, & Allen, 2009; Hambrick & Engle, 2002). This also suggests that under certain circumstances, for items well above span limits, conscious effort in binding may be needed, but overall, the episodic buffer functions without effortful processes. In other words,

> "we now regard [the episodic buffer] as being an important but essentially passive structure on which bindings achieved elsewhere [such as LTM] can be displayed. It remains important in that it allows executive processes to carry out further manipulation. This may in turn lead to further bindings of phrases into integrated sentences…" (Baddeley, 2012, p. 17)

In a similar vein, a recent systematic review of 36 episodic buffer studies suggests that “the experimental approach is still necessary because of the lack of agreement concerning the processes in the EB in terms of its specificity in dealing with different information and the relationship between processes” (de Pontes Nobres et al., 2013, pg. 341). As one can see, much of the work on the buffer seems to still be evolving and changing, with much empirical work still to be done, especially in the realm of how the episodic buffer is involved in L2 language processing.

Chapter 2 - The Current Studies

These data do suggest that the episodic buffer is a psychologically real construct, although it has still not been extensively tested and some previous results have been inconsistent. Furthermore, it is important to note that many of the studies conducted that have examined the processing of prose have focused on recall rather than recognition or comprehension. In addition, to date, there has only been one study that has examined an L2, but it was for sign language (Rudner & Ronnberg, 2008). Furthermore, being a transient system, the episodic buffer can create novel representations, when background knowledge for items like a hockey-playing elephant are not stored in LTM.

Keeping this in mind, it is important to conduct additional research exploring the episodic buffer in several ways. The first is to extend the understanding of the episodic buffer by exploring the reading comprehension abilities of L1 readers with varying levels of domain knowledge of the text. As aforementioned, both WM and domain knowledge predict performance in readers, and corroborative results would help us specifically tap into the episodic buffer especially when text must be comprehended in the midst of a cognitive demanding dual-task. Second is to extend the understanding of the episodic buffer by exploring the reading comprehension abilities of L2 learners of various levels. In other words, by exploring how L2 learners comprehend inferences in different circumstances, we may gain insight into the functioning of the episodic buffer. Given that reading is a complex activity for L2 learners, especially those who do not have a high WM span and may further lack knowledge of specific topics within American mainstream texts. It is assumed the episodic buffer would need to be actively involved in such binding (integrating of information from LTM and the subservient WM stores). This is especially so under cognitively demanding conditions, such as those requiring a speeded response. Specifically, because a cognitively demanding situation for an L2 learner disrupts attention to a task as well as depletes resources, it is possible that under a task that is not particularly difficult, the episodic buffer would automatically be able to bind multimodal information (such as online generation and comprehension of basic facts), whereas in doing the same for a pragmatic inference in which the episodic buffer would need conscious effort to be able to bind information. Thus methodologies with a focus on comprehension and recognition memory rather than recall were explored. Additionally, by providing evidence that L2 learners make novel representations out of text that they may not be familiar with, it would show that the LTWM may not be best suited to conceptualizing reading in an L2 (since LTWM assumes novel representations are not made, but prior knowledge drives performance). As such, the episodic buffer may be a better conceptualization, which is why the episodic buffer is of importance in the current study. Once again, the conceptualization of WM as a construct that includes the episodic buffer allows it to be a more versatile construct in its ability to make several predictions of higher level cognitive processes such as discourse processing, especially when related to stress and inference processes. As such, it seems to be a fruitful area of research in an attempt to understand how the episodic buffer really works, now that we understand that it does indeed exist.

As such, the purpose of the current studies was to ask the questions of how can the study of inference processing under different conditions of varying cognitive demands in L1 and L2 readers contribute to our understanding of the episodic buffer? As aforementioned, some may argue that L2 reading is a declarative process, whereas reading in an L1 using DK is more of a procedural process, and perhaps even use different areas of the brain for those processes.

However, both require the retrieval of information from LTM. As such, they may have enough similarities to be considered for comparison.   In other words, given that a) the reading of text, and further its comprehension, involves the integration of several levels of comprehension, b) the episodic buffer is known as an entity that holds integrated information to be manipulated by CE, and c) that cognitive demands (task difficulty of a secondary task, increasing inferential complexity, domain knowledge) can influence WM, it can be asked how can the study of inference processing under different conditions (varying cognitive demands) in L1 and L2 readers contribute to our understanding of the episodic buffer.

This is important because foreign language reading is not an ‘expert skill’ with respect to language skills, so it should be expected to involve considerable cognitive effort (as indicated by the several studies mentioned above), which is contrary to LTWM, which suggests that reading is indeed an automatic expert skills task. In fact Kintsch (2010, personal correspondence) suggests that the retrieval structures in LTWM become stronger via learning to create a certain level of expertise which makes retrieval from LTWM an automatic process. He further argues that, because reading discourse is assumed to be an expert activity, it is therefore automatic. Because those learning a second language are not yet skilled readers in their L2, their retrieval structures are deficient; thus their reading is not automatic and as such, according to Kintsch, they cannot use the LTWM system. The proponents of the episodic buffer do not make any specific claim about learning and expertise (e.g., language skills) required to access LTM, as it may or may not be a conscious activation of the episodic buffer, although it is well established that those with high WM perform better on language tasks than those with low WM. Thus it is quite possible that L2 learners are able to comprehend discourse using the episodic buffer, but not LTWM. Further, the reading and comprehension of text involves the integration of several levels of comprehension, which is important when thinking about the episodic buffer as an entity that holds integrated information.

If we further consider the conscious effort needed to read prose, then those with low DK of a certain text should be somewhat comparable to those who have less fluent English skills. In other words, domain knowledge and language proficiency should both be proxies for difficulty level. That is, in terms of performance (effectiveness and efficiency) L2 beginners should perform worse than L2 intermediate readers, with L1 English readers showing the best performance, similar to that of the low DK vs. high DK distinction. For L1 readers, reading is a skilled task, but it is effortful for L2 readers, making it harder for the episodic buffer to bind information from LTM. For L2 readers who are not at a highly skilled level in their L2, the episodic buffer should require more effort to bind information, and thus those in lower levels of L2 proficiency should have less efficiency than those with greater L2 proficiency. That is, highly skilled readers (i.e. native English speakers) can use their episodic buffer automatically, so that it mirrors the functioning of the retrieval structures in LTWM. Under circumstances of high cognitive load, they use their episodic buffer with conscious awareness enabling them to comprehend information, thus encompassing more than LTWM. In other words, it is possible that the episodic buffer may be a useful construct for memory performance beyond expert performance alone, wherein in certain circumstances it acts like LTWM, but in other circumstances operates as an independent construct in effortful processing.

Taken together, this suggests that the study of L2 readers' comprehension of inferences can allow insight into how the episodic buffer works, while also allowing for further insight to the many differences that exist between L2 learners and native L1 readers. Again, as mentioned previously, given that a) reading (and comprehension) of text involves the integration of several levels of comprehension b) the episodic buffer is an entity that holds integrated information to be manipulated by the central executive and c) cognitive demands (task difficulty, domain knowledge) can influence WM, I can ask how can the study of inference processing under different conditions (varying cognitive demands) in L1 and L2 readers contribute to our understanding of the episodic buffer.

To answer these questions, Experiment 1 examined how reading comprehension of L1 readers is influenced by increasing cognitive load (dual-task and increased inferential complexity), domain knowledge, and WM. Specifically, by this experiment further insight as to how the episodic buffer functions can be gained. Given that we are using domain knowledge and language proficiency as proxies for difficulty, Experiment 2 will further enable us to understand how differences in L1 and L2 reading comprehension can help us gain further understanding of the episodic buffer. Similar to Experiment 1, reading comprehension in both L1 and L2 readers will be examined by assessing how reading is influenced by increasing cognitive load (dual-task and increased inferential complexity) and WM.

Chapter 3 - Experiment 1

Hypotheses and Possible Outcomes

Based on the literature, there are three possible outcomes that we can expect, each which will be discussed in turn.

The first possible outcome, as seen in Figure 1, would be that the differences between high and low DK become smaller with increasing task difficulty.

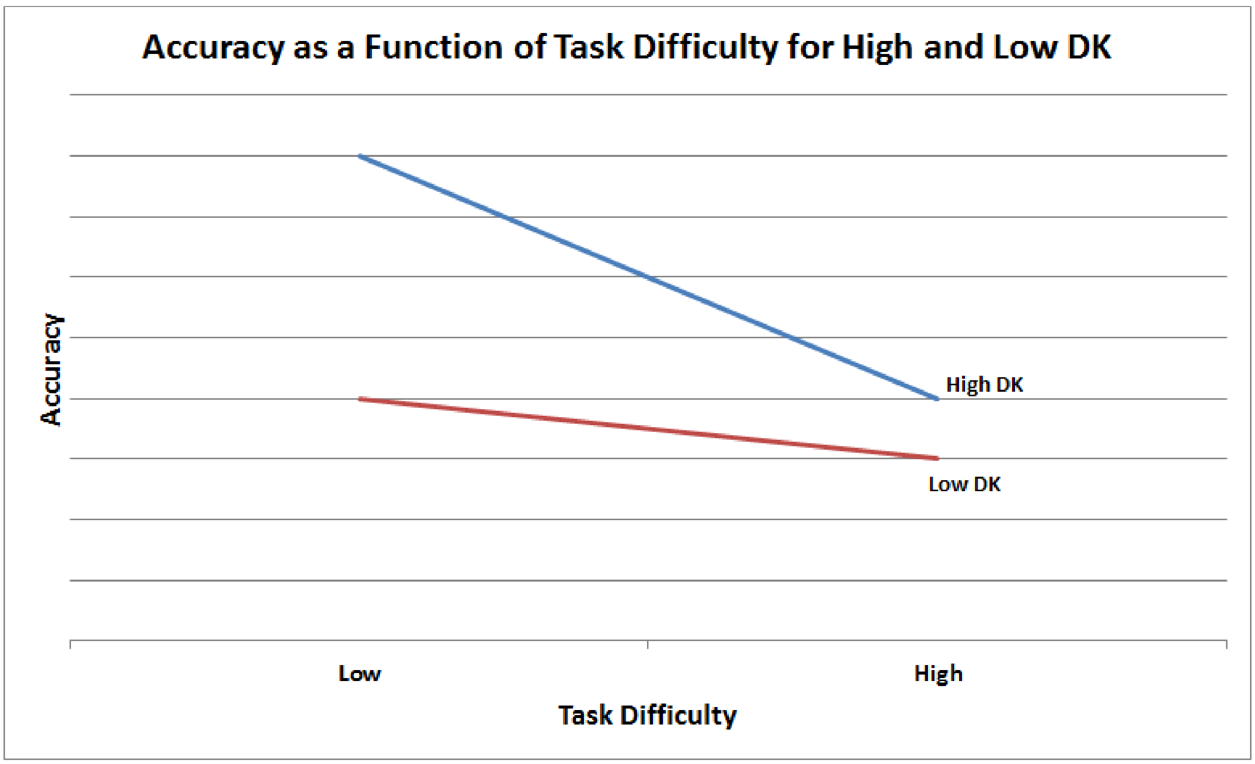

Figure 1. Possible Outcome 1.

That is, as difficulty increases (due to cognitive load, or increasing inferential complexity), performance should decrease based on the assumption that, as the central executive is taxed, it would be more important in performance than the episodic buffer (so conscious awareness would be needed), resulting in poorer performance with greater cognitive load. In other words, the logical (face value) assumption would be that as task difficulty increases, performance decreases. As such, those with higher DK are able to use their episodic buffer to retrieve into LTM (i.e. episodic buffer can get info from LTM automatically). However, if there is extra cognitive load, such as the secondary task, the episodic buffer uses conscious awareness so CE cannot handle the extra load. This extra cognitive load results in a breakdown in relay of information between the CE and episodic buffer, resulting in poorer performance.

The low DK readers are already using most of the resources available in CE, with less DK knowledge to begin with, which is needed to activate the situation model (to obtain info from LTM). That requires low DK readers to use more conscious effort for the basic reading task, so even without added cognitive load they do poorly, and cannot really get any lower (since comprehension performance is near floor effect already). This reasoning would also be a possible outcome for Reaction Time (RT) also in that, as the cognitive load increases the reaction time increases, but more so for those with lower DK. This outcome would be consistent with Hypothesis 1.

H1: Given that the episodic buffer is supposed to work only as well as the central executive can allow it to bind information, with more difficult tasks requiring more resources, and therefore conscious effort from the episodic buffer, readers with adequate domain knowledge (DK) of text (easy text) should be able to comprehend inferences with more effectiveness (accuracy) and efficiency (reaction time) than those with less domain knowledge of the text because those with higher DK should be able to retrieve information from LTM better than those with low DK

The second possible outcome, as seen in Figure 2, is that the difference between those with higher DK and lower DK become greater as the task difficulty increases (when cognitive load is present or inferential complexity increases).

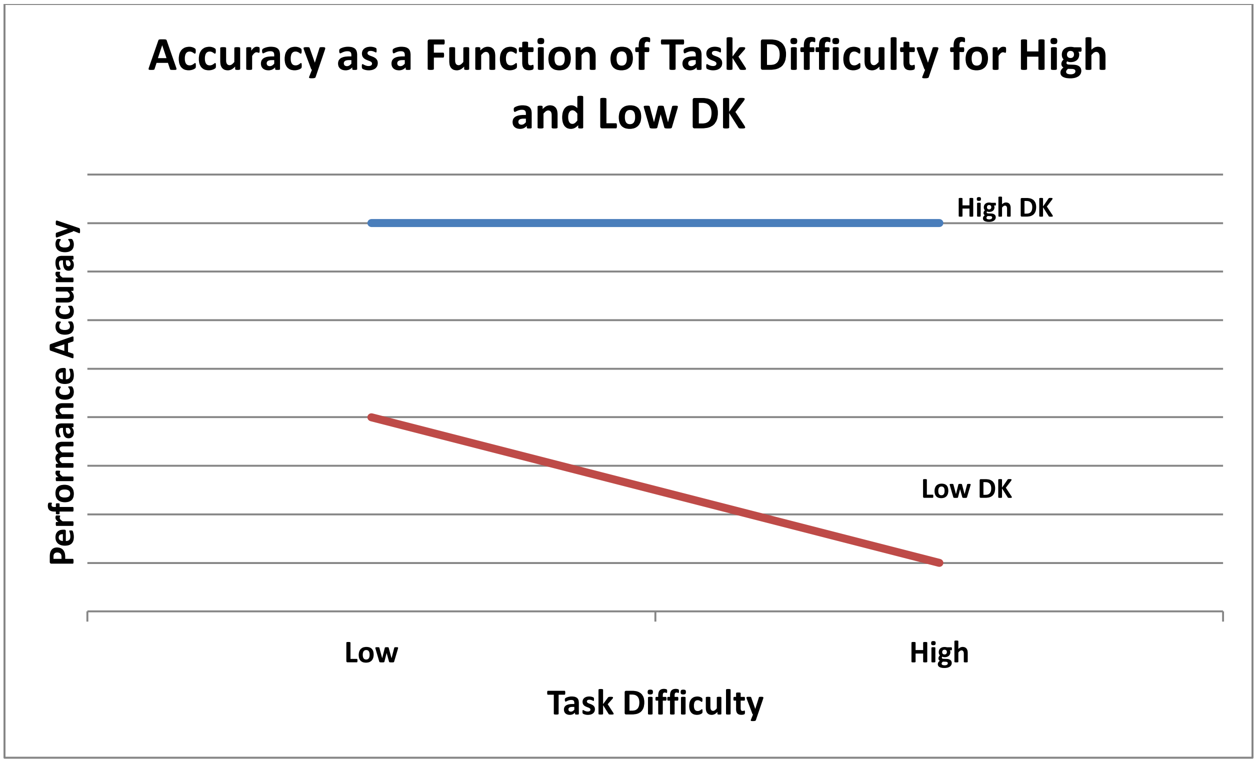

Figure 2. Possible Outcome 2.

This is based on the notion that those with high DK have sufficient resources to retrieve information from LTM, when cognitive load increases. If the episodic buffer is separate from the CE, it will not be affected by the additional load, whereas for low DK, for whom the task requires much more conscious effort, with extra strain on CE, low WM span will decrease performance. This would be consistent with some previous results (e.g., Experiment 3 from Jeffries et al., 2004) , who found no differences in performance while performing a dual task when recalling prose, thus showing that CE and episodic buffer are distinct entities, and that the episodic buffer can bind information automatically thereby preventing performance decrements. This would be consistent with the 'rich-get richer' hypothesis (Hypothesis 2), whereby those with high DK also have enough WM resources to facilitate performance and withstand the effect of the increased cognitive load from a dual task. Similarly, this outcome can be used to explain the possible effects on RT also, where RT increases only for those with Low DK, as they do not have the sufficient resources to retrieve information from LTM, and therefore take longer to bind information together (using conscious effort), whereas those with high DK are able to bind information automatically and therefore take less time to respond. This is consistent with the rich-get richer hypotheses in that high WM span facilitates reading comprehension on tasks that require reading in a first language (Rai et al, in press). Further studies that argue for the rich-get-richer hypothesis (Cantor & Engle, 1993; Just & Carpenter, 1992), high WM should just facilitate high domain knowledge even further, (rich get richer hypothesis) than those with low WM and high DK.

The final possible outcome, as displayed in Figure 3, would be that reflected in two main effects with no interaction between domain knowledge and task difficulty, in that there would be no differences between those with high and low domain knowledge as difficulty increases.

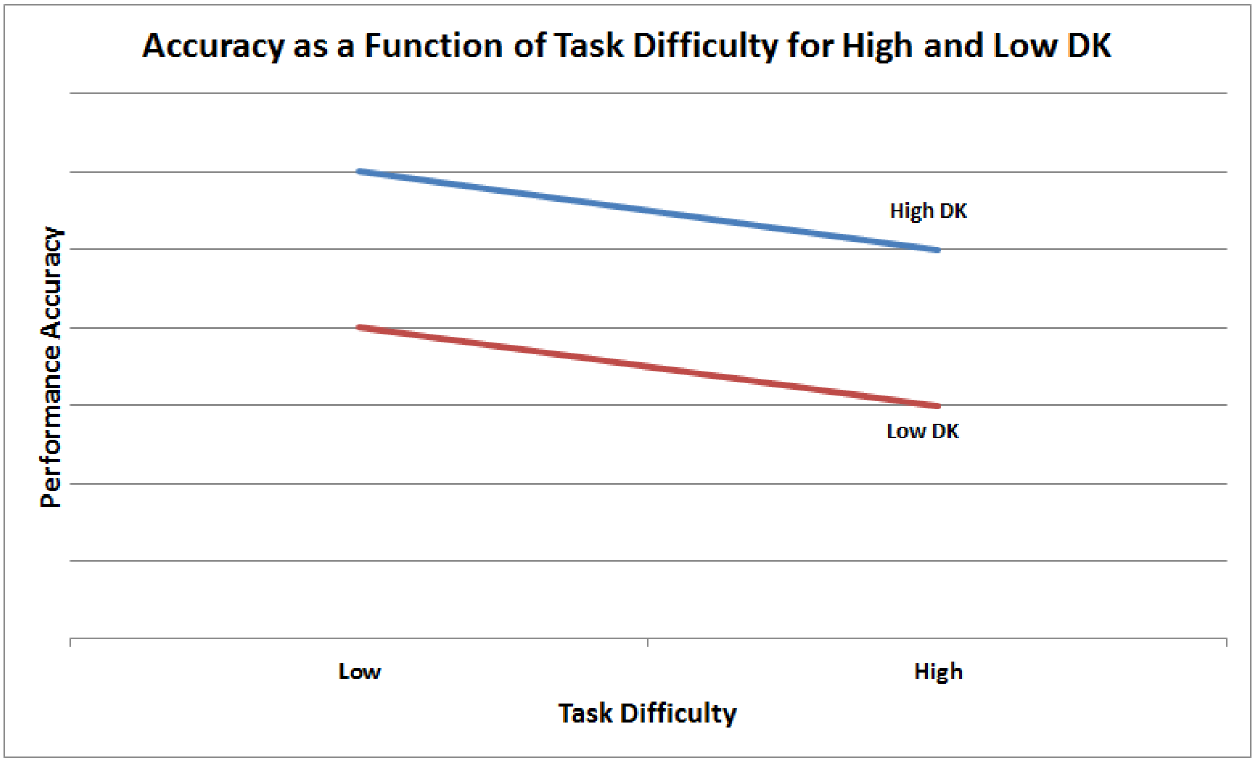

Figure 3. Possible outcome 3.

This would be similar to outcome #1, except that both groups would systematically decrease performance about the same with task difficulty. This would suggest that the central executive is of importance for both groups, both of which are taxed by the cognitive load. This suggests that the episodic buffer is not able to integrate information from LTM to the central executive very well, so the CE in this case matters more in performance than does the episodic buffer. For the reaction time here, some previous studies have indicated that typically as cognitive demands are increased, there is a speed-accuracy tradeoff in performance. Given that this task will be self-paced, it is possible that RT may systematically increase as the task becomes more difficult, without compromising accuracy. Similar to the reasoning behind the systematic decline in effectiveness, the systematic decrease in efficiency suggests that the central executive is important for both groups, in which they both are taxed by the cognitive load, thereby take longer overall. This suggests that the episodic buffer is not able to integrate information from LTM to the central executive very well, so the CE in this case matters more in performance than does the episodic buffer.

Furthermore, regardless of DK, those with higher WM should respond faster than those with lower WM, due to their increased ability to use available cognitive resources when needed to the appropriate task.

In terms of the inference type questions, as the complexity of the inferences increases, effectiveness should decrease, especially for those with lower WM, due to their inability to use their cognitive resources to focus on the situation model. This should be especially so when the complexity of the inference is coupled with the dual-task. Previous studies (Eysenck et al, 2007; Rai et al., 2011) have found detrimental effects on effectiveness as cognitive load (such as stress in those studies) is introduced. Additionally, in terms of the inference type questions, as the complexity of the inferences increases, effectiveness should decrease especially for those with lower WM due to their inability to use their cognitive resources to focus on the situation model. This should be especially so when the complexity of the inference is coupled with the dual-task. Previous studies (Eysenck et al, 2007; Rai et al., 2011) have found detrimental effects on efficiency as cognitive load (such as stress in their case) is introduced. To the degree that episodic buffer is able to use prior knowledge, and the CE is able to integrate that information, leaving resources available for the tasks that have a greater cognitive demand (increased complexity of inferences and dual-task) those with higher WM should perform better.

In the first experiment, Hypothesis 1-3will be addressed with L1 readers only. In Experiment 2, L1 readers will be used as the control group to assess how the episodic buffer works when examining L2 performance.

Method

Participants

Sixty-seven native English speakers from undergraduate psychology courses in a Midwestern University participated in the study. They had a mean age of 19.25, SD = 1.59, with 36 men and 32 women. All of the participants reported English reading proficiency levels which were quite homogenous, with 95.6% rating themselves as good or excellent on a 5-point reading proficiency scale (poor = 1, fair = 0, average = 2, good = 22, excellent = 43). All of the students were native English speakers and thus were included in the analyses.

Measures

Demographic Questionnaire: All participants completed a short demographic questionnaire (Appendix A) developed by the experimenters which included the question of self–rated reading proficiency described above. In addition to demographic questions and reading assessment, this questionnaire also had 5-point assessments of their writing, listening, and speaking skills in English, along with questions about any other languages they might know. For this particular experiment, self-rated reading proficiency was of most importance as it has been found to correlate with actual performance of inference comprehension while reading (Dufour & Kroll, 1995; Kroll et al., 2002; MacIntyre, Noels, & Clément, 1997; Rai et al., 2011). Thus self-rated reading proficiency was used as a predictor in the study.

Automated Operation Span Task (OSpan). This computerized version of the OSpan task developed by Unsworth et al. (2005) is an established measure of working memory. The OSpan task particularly taps into the central executive component of WM by forcing participants to remember a string of letters while performing simple mathematical operations between the presentation of each letter. Because the OSpan measures the capacity of the central executive in

WM, it predicts reading ability without measuring reading ability itself. For example, it is both highly correlated with reading span tasks (.88), which directly involve reading, as well as with mathematical operation span tasks (.77), which do not (Unsworth et al., 2005).

Baseball Knowledge Test: All participants completed Spilich et al's (1979) 45-item questionnaire assessing baseball knowledge, from which 42 were used in the present study (Appendix B). This questionnaire has been widely used and is a well-established test for baseball knowledge. Previous research has indicated that is distinguishes well between those with high vs. low domain knowledge well (Hambrick & Engle, 2002; McNamara & McDaniels, 2004; Ricks & Wiley, 2009).

Reading Comprehension Task (RCT). This is the primary task of stories to be read. The computerized (RCT) was in English to measure readers' comprehension (Appendix C). There were 4 baseball-related stories each presented as a whole, approximately 300-350 words in length. These stories have been previously been used by other studies assessing WM (Hambrick & Engle, 2002; Spilich, 1979) and were slightly adapted to ensure each inference type question would be represented in the text. Two stories were presented in the first block, and two in the second block, with the order of stories counterbalanced across both blocks. Following each passage were nine questions about its content, three for each of three inference types presented in random order for each participant. Three questions were factual questions which involved no inference processes. Three other questions required the participant to draw a bridging inference to identify a pronominal referent. The final three questions required the drawing of a pragmatic inference.

A 7-point likert scale was used to assess correctness and certainty of response, with the correct response on one end of the scale, and the incorrect response on the other end (Appendix

C). The anchor points for all three inference type questions were counterbalanced to ensure that one end was not always the correct answer. These were later recoded to ensure that all the correct answers were on the high end. As such, responses 1-3 were essentially the 'error' responses, 4 being 'I don't know' and '7' being the most correct answer. There was also a practice story at the start of each block of stories (without and with the cognitive load described below). Each story appeared as a whole on the screen, with each question appearing individually. See Appendix A for examples of the story, along with questions and their anchors.

Cognitive Load Manipulation (CL). In the cognitive load condition, participants had to distinguish between a high tone (660 Hz) or a low tone (440Hz) as they were reading the story. These tones were within the range of those used in other similar dual-task studies that had participants distinguish between tones when conducting another primary task (Lufti, 2008; McAnnally et al., 2004). The tones were presented via the speakers of the computer, and presented dichotically through simple headphones. The tones were presented at random intervals of 2, 3, or 6 seconds. The tone continued until the participants responded (by pressing the up or down arrow key for the high and low tones respectively), or for 3 seconds, whichever came first. The tones were only present during the reading of the main passage, and not during the questions. Essentially the cognitive load condition assessed the role of WM in the comprehension of inferences in a dual-task paradigm to assess cognitive functioning. This task puts a strain the central executive by dividing attention needed for the task at hand, which is to read the story well enough to answer questions about it later. Since the secondary task is an auditory task, it should put a strain on the phonological loop, (hence a cognitive load) since reading long prose requires the integration from LTM information such as language structure, vocabulary and content of the passage, it is assumed that the episodic buffer should play a role. Therefore, this task should be a beneficial one to use for both L1 readers and L2 learners (tested in Experiment 2) (since responding to tones is language-independent), who may exert various levels of effort when reading during a dual-task paradigm.

Design and Procedure

Each participant was tested individually, and all items were presented in English during two testing sessions. In the first session (on the online Sona systems for questionnaires), participants filled out the demographic questionnaire and the baseball knowledge test. In the second session (only available for those who had completed session one), participants were tested individually in the psycholinguistics laboratory. First, participants completed the Automated Operation Span task. Participants then completed the RCT task divided into two blocks, with 2 stories in each block. For the first block, participants conducted the RCT task without a secondary task. In the second block, participants were told that they now need to continue to read the stories as they have done for the first block, but now they will hear either a high tone or a low tone and need to respond to that tone by pressing either the up arrow or down arrow key. There was a practice session for distinguishing between the tones, followed by a practice session of reading while responding to the tones. The practice sessions mirrored the actual experiment in terms of length of the stories, but had only 2 questions per inference type instead of three. Participants were told to do their best on answering questions about the stories. In other words, all participants were given practice trials with the same practice story to familiarize themselves with the task at the beginning, and then again before the block where the secondary task was introduced. After they read all four passages and answered the accompanying questions for each passage, they were debriefed.

Recap of Measured Variables

The dependent variables in this study were the certainty scores (a measure of accuracy) of the difference inference type questions in the RCT. Additionally, the reaction time (RT) to answer the different inference type questions was also assessed.

The independent variables were as follows: WM capacity and domain knowledge were both continuous between-subject variables. The type of inference question (non-inference/fact, bridging, or pragmatic) and task difficulty (Cognitive Load present or absent) were both categorical within-subject variables.

Results and Discussion

In order to assess the effects of inference comprehension on accuracy and RT, participant responses were recoded where '7' was always the correct (highest certainty) end of the 7-point likert scale. The responses were recoded because the direction of the correct answers in the questions (as seen in Appendix A) were counterbalanced, where sometimes one anchor was correct, and sometimes the other. As such, recoding ensured that the higher the response, the more correct. As such, responses 1-3 were essentially the 'error' responses, 4 being 'I don't know' and '5-7' being the most correct answer. Additionally, responses were collapsed to create a two-choice correct response. That is, 1-4 were recoded as '0' for incorrect responses, and 5-7 recorded as '1' for correct responses to assess percent correct in the accuracy analyses. For RT analyses, only these correct responses were used.

Accuracy

Before conducting regression and ANOVA analyses on the variables, their descriptive statistics were examined. The means, standard deviations, and range of this exploration are displayed in Table 1. It is important to note, that even though the mean of the BBK seems low, it does have a normal distribution, and there were some that scored quite high out of 42. Generally, the means of all three variables are similar to those of other studies.

Table 1. Mean, Standard Deviation and Range of continuous variables

Variable	Mean (SD)	Possible Range
Working Memory (OSpan)	55.26 (14.31)	9-75
Baseball Knowledge (BBK)	18.09 (12.25)	0-42
Self-Reported Reading Skills	4.58 (.692)	1-5

Additionally in order to assess the relationship between these predictor variables of OSpan, BBK, and SRRS, person correlations were calculated as shows in Table 2.

Table 2. Correlations between continuous variables

	Working Memory (OSpan)	Baseball Knowledge (BBK)	Self-Reported Reading Proficiency (SRRS)
OSpan	1		
BBK	.532*	1	
SRRS	.237	-.040	1

Note: * p < .001

Simultaneous multiple regressions were conducted to further probe the degree to which working memory, baseball knowledge and reading proficiency uniquely explained variance in the different outcome measures for each combination of accuracy condition (3 types of inferences by cognitive load condition: no load vs. high load). Working Memory (OSpan), BBK and Self-reported reading skills (SRRS) were the predictors. Since six separate regressions were conducted, a Bonferroni correction was used to allow for a better representation of the results. This correction changed the p-value be significant from .05 level to a p=.008 level to reach significance. Table 3 summarizes these regression analyses.

Table 3 Summary of simultaneous regression analyses for variables predicting accuracy in combinations of DVs Beta (s.e) from least to most complex.

	Fact (no load)	Fact (with load)	Bridging (no load)	Bridging (with load)	Pragmatic (no load)	Pragmatic (with load)
(OSpan)	.002 (.002)	.001 (.002)	.005 (.002)	.002 (.003)	.004 (.002)	.004 (.003)
(BBK)	-.003 (.003)	.004 (.002)	-.001 (.003)	.004 (.003)	.006 (.003)	.005 (.003)
(SRRS)	.011 (.040)	.001 (.035)	-.009 (.040)	.038 (.049)	-.023 (.042)	.021 (.048)
R^2	-.018	.033	.068	.041	.185*	.116

* $p < .008$

As shown in Table 3, there was only one instance where the regression reached the corrected bonferroni level of significance. That is under the pragmatic no load condition, where the continuous variables predicted 18.5% of the variance in the pragmatic no load condition ($R^2 = .185$, p = .001). Upon closer inspection however, it was found that BBK's unique effect on pragmatic no load condition did not reach the p<.008 level, but was approaching significance, *B*=.006, *t* =2.143, *p* = .036. Similarly, although none of the variables predicted any variance in the pragmatic with load condition, the R^2 value was approaching significance, with a p=.015.

Taken together, these suggest that it is only in the most demanding conditions (such as those with pragmatic inferences) that these variables are of most importance. Specifically, domain knowledge is useful to help comprehension; however, it only predicted accuracy for making pragmatic inferences with no cognitive load. Although working memory (OSpan) did not predict any variance in accuracy, it was correlated with BBK, and as such, perhaps the underlying mechanism that drives the DK may be those greater resources in WM, although this claim may be examined in future studies.

In addition to the regressions, a 3 (Inference Type) x 2 (Cognitive Load condition) repeated measures ANOVA was conducted on accuracy. There was a moderate main effect of inference type F (2, 132) = 7.15, p =.001 where accuracy was significantly better for fact (M=.663, SD = .218) and bridging questions (M=.679, SD = .221) than for the pragmatic inference (M=.587, SD = .275) as seen by Bonferroni t-tests $p<.05$, Cohen's f = .25 This is not surprising in that one would expect the greatest decrement in accuracy for the most difficult inference type (pragmatic inferences). Additionally, there was a moderate, (Cohen's f = .238) significant main effect of cognitive load F (1,66) = 8.56, p =.005), with accuracy greater with no cognitive load (M = .675 SD = .245) than with cognitive load (M= .611, SD = .241). There was no interaction between IType and Cognitive Load ($F<1$). Nevertheless, it may be of note to point out that the trend toward decreasing accuracy especially for the most difficult conditions. Generally, accuracy is decreasing from fact to pragmatic questions, especially in the load condition (with the lowest accuracy being for the pragmatic load condition (M=.552, SD = .227). Again, this interaction was not significant, but is trending toward the suggestion made earlier that increasing difficulty leads to greater decrements in performance.

Reaction Time

To further explore the relationship of the working memory and domain knowledge (as well as the other continuous variables), the same simultaneous regressions described for accuracy data above were also conducted on reaction time (RT) data. All RT analyses were conducted only on the correct responses and represent the time it took to both read and answer the questions. Furthermore, outliers (the top and bottom 1%) were trimmed based on z-scores conditionalized on each combination of variables: 2 (Cognitive Load condition) x 3 (Inference Type). This resulted in the removal of 2% of the data from 1567 data points, to 1543 data points left in the analysis. Table 4 summarizes the regressions on RT. As with accuracy, the bonferroni correction made a cut-off for significance at p=.008.

Table 4. Summary of simultaneous regression analyses for variables predicting reaction time in combinations of DVs Beta (s.e) from least to most complex.

	Fact (no load)	Fact (with load)	Bridging (no load)	Bridging (with load)	Pragmatic (no load)	Pragmatic (with load)
(OSpan)	-.337 (16.36)	7.31 (13.98)	7.04 (17.05)	3.77 (13.20)	6.70 (22.52)	18.33 (19.20)
(BBK)	-41.14 (19.17)	-22.52 (15.76)	-14.22 (19.89)	-8.11 (15.15)	18.61 (26.17)	-5.49 (21.63)
(SRRS)	-444.78 (284.04)	88.50 (232.61)	-58.74 (298.34)	-34.87 (222.10)	-412.20 (393.78)	-168.30 (332.09)
R^2	.089	-.008	-.040	-.044	-.012	-.034

Of the six regressions, none reached the .008 level as the continuous variables did not predict in variance in any combinations of the dependent variables. Given that BBK predicted variance in only the more difficult pragmatic inference, it may be possible that domain knowledge is indeed very important especially at the later stages of reading comprehension, but that RT is not impacted as that difficulty increases. It may also mean that resources and DK are not needed at the early stages of reading comprehension, such as those requiring only text-based levels of comprehension, and thus do not impede the central executive as much resulting in less of a need for the episodic buffer (as readers can go faster with more domain knowledge, which then does not interrupt their accuracy).

To further probe the effects on RT, a 3 (Inference Type) x 2 (Cognitive Load condition) repeated measures ANOVA was conducted on RT. There was a statistically significant main effect of Inference Type on RT, F (2, 118) = 43.73, $p < .001$, with RT increasing as inferential complexity increased from fact (M = 4951 msec, SD = 1539 msec) to bridging inference (M =

5117 msec, *SD* = 1552 msec) to pragmatic inference (*M* = 6353 msec, *SD* = 2028 msec). Bonferroni pairwise comparisons revealed a significant difference between pragmatic inferences and fact and bridging inferences ($p<.001$), with the latter two not being significantly different from each other, Cohen's $f = .778$.

Additionally, there was a statistically significant main effect of cognitive load on reaction time, $F(1, 59) = 73.65$, $p < .001$. Cohen's $f = .561$. Surprisingly, RT was greater when there was no cognitive load (*M* = 6105 msec, *SD* = 1881 msec) than when a cognitive load was present (*M* = 4842 msec, *SD* = 1489 msec). Although the Cognitive Load (No load vs. load) by Inference Type interaction was not significant ($F < 1$), there was a trend toward the greatest differences in RT between load and no load condition for pragmatic inferences more than the fact and bridging inference ones. This suggests that as complexity is increases, it does have a greater impact on performance, in this case, with RTs decreasing.

It is interesting to note that, although RT generally increased as inferential complexity increased for both cognitive load conditions, when cognitive load is present, participants were actually responding faster than when no load was present. The fact that people are going *faster* under cognitive load suggests that the perhaps the episodic buffer is indeed playing an active role in inferential processing here, as the CE is not showing detrimental effects of the greater load but in fact advantageous ones. If taken together with the accuracy data where the only significant regression was at the most complex pragmatic inference, perhaps readers are able to maintain accuracy while keeping RT intact because of the active role of the episodic buffer. Perhaps also, since the more complex inferences, namely the pragmatic inferences, are the ones where more domain knowledge may be needed as answers are required from baseball knowledge as a whole rather than what is in the text alone, the episodic buffer is facilitating the comprehension of domain knowledge, thus resulting in lower RTs in the high cognitive load conditions. To further explore this relationship, the accuracy and RT means were compared under high and low cognitive load conditions to understand if there may be any speed accuracy tradeoffs occurring. With no load present, the mean accuracy is .675 with a mean RT of 6105 msec. Under conditions of load, both accuracy and RT decreased to .611 and 4842 msec. This suggests that there is no speed accuracy tradeoff, but as suggested earlier, perhaps the episodic buffer is able to facilitate some of the load, and helps decrease RT. It is possible if participants spent more time responding to the questions under the cognitive load condition, then perhaps accuracy would not have decreased. Additionally, future studies will implement the linear mixed modeling approach to directly compare the three outcomes with the current data, although it does seem as though there is a mix of outcome 2 and 3 occurring here, since there are no interactions, but still at the most difficult levels of task difficulty, there seem to be effects of DK, and possibly of RT (due to the decreased RTs). As aforementioned, the linear mixed modeling approach to be used in future studies may help parse out some of these details.

Chapter 4 - Experiment 2

As mentioned above, if we consider conscious effort needed to read prose, then those with low DK should be in a similar position to those who have less than fluent English skills. In other words, DK and language proficiency should be proxies for difficulty level. It is this variable of language proficiency (L1, and 2 levels of L2) which is the basis of Experiment 2 to compare L1 and L2 readers.

Since we are considering L2 proficiency and DK as a proxy for difficulty level, the hypotheses for Experiment 2 coincide closely with that of Experiment 1.

Given that we assume that L2 proficiency and domain knowledge are a proxy for task difficulty, the same 3 possible outcomes that may come about from Experiment 1 would be similar to expect for Experiment 2, wherein native English readers, those with higher L2 proficiency would be similar to the high DK group, and those with lower L2 proficiency being similar to those with low DK. It is important to note that similar results as Experiment 1 do not mean that language proficiency is qualitatively the same or comparable to domain knowledge. However, it may mean that both use the same underlying processes of the episodic buffer, and therefore are indeed a proxy for task difficulty.

Method

Participants

This study had 3 Groups divided based on proficiency levels. The first group consisted of 72 Native English speakers (all listing themselves as 'American') from the general psychology participant pool at a Midwestern U.S. university. Group 1 (Native English Readers) had a mean age of 18.9, SD = 1.41, with 28 men and 48 women, and 2 undisclosed. Group 1 participants reported English reading proficiency levels, which were quite homogeneous, with 97.2% of the participants rating themselves as average, good or excellent on a 5-point reading proficiency scale (poor = 0, fair =2, average = 5, good = 23, excellent = 42) (M = 4.46, SD = .749).

The second group (Intermediate English Readers) consisted of 62 intermediate English proficiency students. These students were all enrolled in the English Language Program (ELP 152) reading classes based on internal placement test scores that placed them in the intermediate reading classes. The mean age of group 2 was 25.08, SD = 5.28, with 30 men, 29 women, and 3 undisclosed. Although all of the students were in the same class, their nationalities were Chinese (N=31), Ecuadorian (N=23), Saudi (N=6), Iranian (N=2), and one each Indian and Paraguayan. This was further reflected in Group 2's native language: Chinese (N=28) Spanish (N=24), Arabic (N=6), Persian (N=2) and Telegu (N=1). Additionally, Group 2 was quite homogenous in their self-reported English reading proficiency levels, with 90% of participants rating themselves as average or good on a 5-point reading proficiency scale (poor = 1, fair = 3, average = 32, good = 22, excellent = 2) (M = 3.35, SD = .709).

The third group (Beginning English Readers) consisted of 38 beginning level English reading proficiency students. These students were all enrolled in the ELP 150 reading classes based on internal placement tests scores that placed them in this particular reading class (where they are not in the very first ELP reading class, but still beginners, although able to do the tasks at hand, as suggested by the ELP program directors). Group 3 had a mean age of 23.61, SD = 6.58, with 23 men, 14 women, and 1 undisclosed. The nationality of the students was primarily Chinese (N=19), with several Spanish/Portuguese speaking backgrounds (Ecuadorian N=5, Brazilian N=4, Bolivian N=1), a few from the Middle East (Saudi N=4, Kuwaiti N=3, Iraqi N=1) and 1 Japanese participant. This was further reflected in Group 3's self-reported native language:

Chinese (N=18), Spanish (N=6), Portuguese (N=5), Arabic (N=8) and Japanese (N=1). Additionally, Group 3 was quite homogenous in their self-reported English reading proficiency levels, with 81.6% of participants rating themselves as average or good on a 5-point scale reading proficiency scale (poor = 1, fair = 5, average = 19, good = 12, excellent = 1) (*M* = 3.18, *SD* = .800).

Both the Intermediate and Beginner English readers received extra credit at the discretion of their teachers for their participation in the study.

These latter two L2 groups were reassigned based on their lexical decision task (LDT) scores - a measure of reading proficiency (discussed in more detail later). This kept the native English readers group as is. The two other groups were divided into low and high LDT scores (resulting in 40 participants per group), based on a median split of the LDT scores. The high LDT (intermediate English readers) Group was quite homogenous in their self-reported English reading proficiency levels, with 90% of participants rating themselves as average or good on a 5-point scale reading proficiency scale (poor = 0, fair = 3, average = 20, good = 16 excellent = 1) (*M* = 3.38, *SD* = .667). The low LDT (beginner English readers) Group was also was quite homogenous in their self-reported English reading proficiency levels, with 77.5 of participants rating themselves as average or good on a 5-point scale reading proficiency scale (poor = 2, fair = 5, average = 17, good = 14 excellent = 2) (*M* = 3.23, *SD* = .920). It is with these latter two rearranged proficiency groups that further analyses were done for Experiment 2.

Measures

Demographic Questionnaire. All participants completed a short demographic questionnaire (Appendix D) developed by the experimenters which included questions of self – rated reading proficiency as described above. In addition to demographic questions, this questionnaire also had 5-point assessments of their writing, listening, and speaking skills in English, along with questions about their daily use of the English language (i.e. hours spent in the language lab doing English language homework etc). As in Experiment 1, for this particular experiment, self-rated reading proficiency was of most importance as it has been found to correlate with actual performance of inference comprehension while reading (Dufour & Kroll, 1995; Kroll et al., 2002; MacIntyre, Noels, & Clément, 1997; Rai et al., 2011) which is one of the variables that was tested.

Automated Operation Span Task (OSpan). The same task from Experiment 1 was used. It is important to note once again that because the OSpan measures the capacity of the central executive in WM, it predicts reading ability without measuring reading ability itself. For example, it is both highly correlated with reading span tasks (r =.88), which directly involve reading, as well as with mathematical operation span tasks (r =.77), which do not (Unsworth et al., 2005). In the same way, the OSpan task seems relatively logically independent of foreign language proficiency. This is consistent with the idea that each person has an underlying language-independent WM span regardless of language, as shown by Van den Noort, et al.'s (2006) significant correlations between WM spans across learners' L1, L2, and L3. However, we would also predict, consistent with Van den Noort, et al.'s (2006) arguments, that performing a reading task in one's L2 will add a processing load that taxes one's executive WM resources. In sum, the "language-free" OSpan executive WM measure is based on the assumption that executive WM capacity and L2 proficiency are separable, but that they interact, such that those lower in L2 proficiency will put a greater strain on their existing central executive WM resources.

Lexical Decision Task (LDT). To have another supplemental measure of reading proficiency, all participants completed the well-established lexical decision task from the

“LanguageMAP.org” site of Harrington, Ingram & Proctor (2014). This online task takes about 15 minutes as the participant responds to two LDT tasks, with four levels of words and pronounceable non-words, by pressing either the left or right arrow keys to indicate whether the letter string is a word or not. Each Test (Test A and Test B) consists of 76 items, of which 56 are words and 20 are nonwords, all controlled for word length. The 56 words consist of 14 items each from four levels of frequency bands (from most to least frequent): level 1 (e.g., *still, key, sheet, student*); level 2 (e.g., *cable, mechanic, sweep*); level 3 (e.g., *abide, calf, retrospect*) and level 4 (e.g., *polymer, hew, bewitch*). Examples of non-words are *berrow, mork,* and *charp*. The order of items presented was randomized for each participant. Scores were computed for both accuracy and RT. Accuracy was converted into d’ scores, calculated as the Zscore of Hits – Zscore of False Alarms. Participants in the study who had false alarm rates of greater than 50% were excluded from the study, resulting in a total of 68 participants in Group 1 (Native English Readers), and 40 each in the low and high LDT Groups. Reaction time LDT data for correct responses only (i.e., hits) was also trimmed and used in the RT analyses, as other typical RT studies.

Reading Comprehension Task (RCT). Similar to Experiment 1, the primary task was stories to read followed by questions to respond to (Appendix E). The self-paced computerized reading comprehension task (RCT) measured readers’ comprehension in English. There were 6 stories each of similar length about 300 words each. All of the stories were short fictional stories that have been used in previous studies (Rai et al, 2011) to assess reading. Unlike those used in Experiment 1, none of the stories required any specialized domain knowledge. Following each passage were nine questions about its content, three for each of three inference types presented in random order for each participant. Three questions were factual questions which involve no inference processes. Three other questions required the participant to draw a bridging inference to identify a pronominal referent. The final three questions required the drawing of a pragmatic inference. These latter questions required processing at the situation model level, while the factual and bridging inference questions can occur completely within the textbase level of representation (Kintsch, 1998), although the pronominal reference questions required connecting material across two adjacent sentences. A 7-point likert scale was used to assess correctness and certainty of response, with the correct response on one end of the scale, and the incorrect response on the other end (Appendix E). Immediately after reading each passage, participants answered nine questions: three non-inference (factual) questions, three bridging inference (pronoun-referent) question and three pragmatic inference questions, with the order of the six questions randomized for each participant for each story.

For the bridging inference questions, the distance between pronouns and their referents varied from 0-2 sentences, the typical range of distances found in natural texts (Hobbs, 1977). As in Experiment 1, the RCT task was self-paced with one story shown per screen and participants pressing the space bar to proceed to the questions once they read the entire story. Each story was presented as a whole on the screen with each question presented individually one at a time per screen.

Cognitive Load Condition (CL). As outlined in Experiment 1, the same manipulation was used in Experiment 2.

Design and Procedure

Each participant was tested individually in a single session, and all materials were presented in English. First, participants completed the demographic questionnaire, then the LDT task followed by the Automated Operation Span task. After the Automated OSpan task, participants took a 5-minute break before completing the last part of the experiment. Participants then completed the RCT task divided into two blocks (without the cognitive load and with cognitive load), with 3 stories in each block. For the first block, participants read 3 stories and performed the RCT task with no cognitive load. In the second block, participants were told that they now would need to continue to read the stories as they had done for the first block, but now they would hear either a high tone or a low tone and need to respond to that tone by pressing either the up arrow or down arrow. There was a practice session for distinguishing between the tones, followed by a practice session of reading while responding to the tones. However, they were told to do their best on answering the questions on the stories. In other words, all participants were given practice trials with a practice story to familiarize themselves with the task at the beginning, and then again before the block where the secondary task was introduced. After they read all the passages and answered the accompanying questions for each passage, they were debriefed.

Recap of Measured Variables

The dependent variables in this study were the certainty scores (a measure of accuracy) to the different inference type questions in the RCT. Additionally, the reaction time (RT) to answer the different inference type questions was also assessed.

The independent variables were as follows: Working Memory (as measured by OSpan), SRRS, and LDT were all continuous between-subject variables. The type of inference question (non-inference/fact, bridging, or pragmatic) and task difficulty (Cognitive Load present or absent) were both categorical within-subject variables. Group (3 proficiency levels) was the final categorical between-subject variable.

Results and Discussion

Data Preparation

Exactly as in Experiment 1, participant responses were recoded to make ‘7’ always be the correct end of the 7-point likert scale. The responses were recoded because the correct answers to the questions (as seen in Appendix E) were counterbalanced, where sometimes one anchor was correct, and sometimes the other. As such, recoding ensured that the ‘7’ end was always the correct anchor. Additionally, this scale was broken down into the dichotomous variable of correct (5-7) and incorrect (1-4). As in Experiment 1, this correct response was used in all accuracy analyses, with the correct responses only used in the RT analyses.

Accuracy

Before conducted regression and ANOVA analyses on the variables, their descriptive statistics were explored. The mean, standard deviations, and range of this exploration are displayed in Table 5. It is worth pointing out that, although the self-reported reading skills for the intermediate and beginner English reading groups are similar, the beginner group may have been responding based on their specific class (i.e. comparing against their peers) versus their English skills in general. Thus, this may have led to higher ratings in their self-reports because they were comparing among their peers instead of their English skills in general. This possibility must be taken into account in considering SRRS data across the two L2 groups. Also, a reminder that the intermediate group is referring to the high LDT L2 Group, and the beginner group is referring to the low LDT L2 Group.

Table 5. Mean, Standard Deviation and Range of Continuous Variables

Variable	Range	English Readers Group	Mean (SD)
Working Memory (OSpan)	9-75	Native	54.12 (13.46)
		Intermediate	51.95 (17.17)
		Beginner	55.38 (14.11)
Self-Reported Reading Skills (SRRS)	1-5	Native	4.41 (.814)
		Intermediate	3.38 (.667)
		Beginner	3.23 (.920)
Lexical Decision Task (LDT) Zhits-Zfalse alarms	-3.60-2.79	Native	1.19 (.809)
		Intermediate	-.143 (.719)
		Beginner	-1.89 (.513)

Additionally, the Pearson correlations between the continuous variables for all three groups were calculated as seen in Table 6.

Table 6. Correlations between continuous variables for all three groups

	Working Memory (OSpan)	Lexical Decision Task (LDT)	Self-Reported Reading Proficiency (SRRS)
OSpan	1	-.101	-.342*
LDT	-.335*	1	.087
SRSS	-.028	.204	1
OSpan	1		
LDT	.390**	1	
SRRS	.115	.182	1

Note: Coefficients above the main diagonal represent values for the intermediate (LDT high) English readers and the ones below are for the beginner (LDT low) English readers; the last three rows and columns reflect the native English readers. * p<.05, **p<.01

Simultaneous multiple regressions were conducted to further probe the degree to which working memory, lexical decision task scores, and reading proficiency uniquely explained variance in the different outcome measures for each combination of accuracy condition (3 types of inferences by cognitive load condition: no load vs. high load). Working Memory (OSpan), LDT and Self-reported reading skills (SRRS) were the predictors. Since six separate regressions were conducted, a Bonferroni correction was used to allow for a better representation of the results. This correction changed the p-value be significant from .05 level to a p=.008 level to reach significance. Table 7 summarizes these regression analyses.

Table 7. Summary of Regression Analyses for Variables Predicting Accuracy in Combinations of DVs Beta (se) from Least to Most Complex.

	English Readers Group	Fact (no load)	Fact (with load)	Bridging (no load)	Bridging (with load)	Pragmatic (no load)	Pragmatic (with load)
OSpan	Native	.033 (.021)	.025 (.026)	-.002 (.018)	.015 (.029)	.026 (.027)	.073 (.027)
	Intermediate	-.010 (.029)	.032 (.031)	.003 (.032)	.018 (.034)	-.010 (.032)	.046 (.035)
	Beginner	.103* (.036)	.067 (.035)	.058 (.040)	.076 (.034)	.035 (.039)	0.00 (.044)
LDT	Native	.015 (.024)	.002 (.029)	-.006 (.020)	.023 (.034)	.002 (.031)	.099 (.054)
	Intermediate	-.094 (.044)	.034 (.048)	.010 (.049)	-.013 (.052)	.070 (.049)	.099 (.054)
	Beginner	.185 (.070)	.174 (.066)	.071 (.077)	.088 (.066)	.132 (.075)	.051 (.084)
SRRS	Native	.002 (.022)	.001 (.027)	.016 (.018)	-.020 (.030)	.007 (.028)	.013 (.031)
	Intermediate	.078 (.050)	-.033 (.054)	-.060 (.056)	.031 (.059)	.019 (.056)	.018 (.061)
	Beginner	-.075 (.036)	-.044 (.034)	-.043 (.040)	-.072 (.034)	.017 (.039)	.030 (.043)

*p < .008

For the beginner English group, 22% of the variance in the fact-no load condition was predicted by the 3 continuous variables ($R^2 = .22$, p =.007). Specifically, after controlling LDT and SRRS, OSpan uniquely predicted variance in the fact no load condition, B = .103, t = 2.841, p =.007 (Table 7). This suggests that for the lowest proficiency group, WM resources are being utilized at even the easiest levels of comprehension, whereas for the other two groups, they may have enough resources to do the task well without taxing their resources early on. It may be important to point out, that although not significant (missing the Bonferroni correction cut-off, there was a trend towards variance predicted in the most difficult levels of the task (pragmatic load condition) for the Native English readers. That is, OSpan was trending toward having a significant impact in accuracy for the pragmatic no load condition (R^2 = .111, B = .073, t = 2.662, p = .010). Again, it is important to understand that this was not significant, but the trend towards it suggests that one possibility may be of the episodic buffer enabling the native readers to overcome some of the cognitive load by binding the difficult information, especially at the most complex levels of comprehension. In other words, working memory plays a key role in performance. It is perhaps the case that for group 1, as the difficulty increases, the episodic buffer helps to maintain accurate performance. For the beginner English readers, since they are the lower proficiency levels, it is not surprising then that they would need to use all of their resources even at the easiest levels of inferential complexity, as for this group it may be more difficult already, and as such do not have sufficient working memory resources to be available for the more difficult questions, which may explain why OSpan did not predict any variance in accuracy for harder levels of inferential complexity (bridging and pragmatic inferences) for the beginning English readers group.

In addition to these regressions, a 3 (Inference Type (IType)) x 3 (Group) x 2 (cognitive load) repeated measures ANOVA on accuracy was conducted. There was a large effect of group on accuracy F (2, 145) = 33.61, p <.001, Cohen's *f* = .737, as those with the highest proficiency in English, group 1 (*M*=.808, *SD* = .178), were more accurate than group 2 intermediate readers (*M*=.679, *SD* = .232) and group 3 beginner readers (*M* = .632, *SD* = .234). Bonferroni pairwise comparisons revealed that these differences were significant for groups 1 and 2, and 1 and 3 (p<05). This is not surprising, as accuracy was the greatest for those with greater English reading proficiency, with the native English speakers revealing the greatest difference in accuracy from the two L2 groups, and groups 2 and 3 not being that different from each other in accuracy.

There was also a main effect of inference type on accuracy F (2, 290) = 38.21, p<.001, Cohen's f = .289, with decreasing accuracy from fact questions (M = .773, SD = .208) to bridging (M = .711, SD =.205) to pragmatic questions (M = .636, SD =.238). Bonferroni tests revealed that this difference was significant between each inference type from the other (p<.001). There was also a main effect of cognitive load with accuracy greater with no load (M = .732, SD = .217) than with load (M = .681, SD =.228), F (1, 145) = 17.61, p <.001, Cohen's f = .137.

Although not, significant, it is worth pointing out that the interaction between IType and Group was approaching significance, F (4, 290) = 2.07, p = .085. The trend suggests that accuracy was the lowest for the lowest proficiency group at each level of the inference type, with the Native readers being more effective at each level of the inference type (compared to their less proficient Group 2 and 3 counterparts This once again suggests that inference types have a greater impact for those with lower proficiency, especially at the hardest level of comprehension, (namely pragmatic inferences). This also suggests that although there are clear differences between native readers and the other two groups, these two L2 groups may not be as different from each other as initially expected. This issue will be discussed further in the general discussion. Similar to Experiment 1, it may also be worth pointing out that although not significant, the interaction between IType and Cognitive Load approached significance (2, 290) = 2.50, p = .084. Again, although non-significant, the trend is suggesting decreases in accuracy as inferential complexity increases, especially under conditions of cognitive load.

Reaction Time

To further explore the relationship of the continuous variables for the three groups, the same analyses for accuracy data above were also conducted on reaction time (RT) data. As in Experiment 1, all RT analyses were conducted only on the correct responses (5-7) and represent the time it took to both read and answer the questions. Furthermore, outliers (the top and bottom 1%) were trimmed based on z-scores conditionalized on each combination of variables: 2 (Cognitive Load condition) x 3 (Inference Type) x 3 (Group). This resulted in the removal of 2% of the data from 3889 data points, to 3826 data points left in the analysis. Table 8 summarizes the regressions on RT. As with accuracy, the Bonferroni correction made a cut-off for significance at $p=.008$.

Table 8. Summary of Regression Analyses for Variables Predicting Reaction Time for Three Groups in Combinations of DVs Beta (se) from Least to Most Complex

	English Readers Group	Fact (no load)	Fact (with load)	Bridging (no load)	Bridging (with load)	Pragmatic (no load)	Pragmatic (with load)
OSpan	Native	-21.35 (15.44)	-10.06 (11.14)	-12.71 (14.40)	3.21 (10.43)	-18.48 (13.61)	-4.28 (11.32)
	Intermediate	-63.85 (48.97)	-5.90 (20.76)	-10.89 (45.02)	-39.56 (39.15)	-10.30 (47.36)	36.02 (25.72)
	Beginner	-75.32* (26.83)	-2.54 (21.34)	-47.07 (38.77)	1.64 (32.10)	-32.83 (46.67)	13.65 (28.29)
LDT	Native	.222 (1.78)	-.094 (1.28)	.050 (1.66)	-.538 (1.20)	.012 (1.57)	1.18 (1.31)
	Intermediate	-2.79 (4.80)	-.359 (2.04)	-6.55 (4.41)	-.588 (3.84)	-2.99 (4.51)	-.013 (2.52)
	Beginner	3.48 (2.11)	2.23 (1.68)	3.78 (3.05)	-1.69 (2.61)	-.190 (3.65)	.615 (2.21)
SRRS	Native	-580.61 (243.99)	1.88 (176.05)	-374.0 (227.65)	-346.97 (164.81)	-380.46 (215.13)	-263.50 (178.92)
	Intermediate	1249.03 (1286.61)	-522.59 (545.45)	1822.34 (1182.84)	-670.56 (1028.59)	1681.61 (1237.28)	1072.76 (675.68)
	Beginner	350.35 (400.27)	476.17 (318.43)	437.74 (578.50)	408.55 (477.84)	440.25 (694.26)	81.41 (418.42)

*p < .008

As seen in Table 7, only Group 3 (beginner readers) produced any significant results in the regressions. This result however should be interpreted with caution, as the regression itself did not meet to .008 cutoff, although OSpan did. That is, for Group 3 in the fact no load condition, the combination of continuous variables was reaching significance in predicted RT in the fact no load condition (R^2 = .186, p = .022). When controlling for LDT, and SRRS, OSpan uniquely predicted variance in RT for the fact no load condition (B = -75.32, t = -2.81, p = .008) suggested that those with greater WM went faster in the no load condition. This is consistent with the accuracy results that also only had one instance of significance, which was for Group 3. This again suggests that this less proficient group is using up all of their resources earlier on in comprehension, whereas for the other groups, perhaps that task is not taxing enough to account for any significant results.

For accuracy, WM seems to play a key role for the more proficient groups at greater levels of difficulty (see Table 6). It is possible that the episodic buffer is more important for the least proficient group at the easier stages of comprehension, but the buffer continues to monitor performance in terms of RT as the task gets more difficult which is especially prevalent for the least proficient group. Since all of the groups are responding very well (not a ceiling effect, but doing well as seen from the main effect of group on accuracy), perhaps the two most proficient (native and Intermediate reader) groups did not find the need to respond faster, as the task in of itself was not presented as a timed test.

To further probe the effects on RT, a 3 (Inference Type) x 3 (Group) x 2 (Cognitive Load condition) repeated measures ANOVA was conducted on RT. Not surprisingly, there was a statistically significant main effect of Group, F (1, 135) = 54.61, p <.001, Cohen's *f* = .509, with Group 1, native English readers, responding fastest (M= 4593 msec, SD=1446 msec) with RT increasing for Group 2 (M = 8434 msec, SD = 4056 msec) and slightly decreasing again for Group 3 (M = 6278 msec, SD = 2802 msec). Bonferroni pairwise analyses revealed that each Group was significantly different from each other (1vs. 2 and 3, 2 vs. 3) at p < .05. Additionally, there was a main effect of Inference Type *F* (2, 270) = 39.16, *p*<.001, Cohen's *f* = .175. Fact questions were responded to the fastest (*M* =5652 msec, *SD* =2838 msec) with RT increasing from bridging (M = 6690, msec SD=3284 msec) to pragmatic inference (M = 6961 msec, SD = 3159 msec). Bonferroni analyses showed that RT for fact questions was significantly different than bridging and inference questions, with $p<.05$. There was also a main effect of cognitive load $F\ (1, 135) = 70.21$, $p < .001$, Cohen's $f = .204$. Surprisingly, as was found in Experiment 1, RT was faster with cognitive load present (M =5636 msec, SD = 2435 msec) than with no load (M=7234 msec, SD = 3570 msec).

These main effects were qualified by three interactions. First as seen in Figure 4 there was a statistically significant Group x Inference Type interaction $F\ (4,270) = 4.10$, $p<.001$, Cohen's $f = .100$

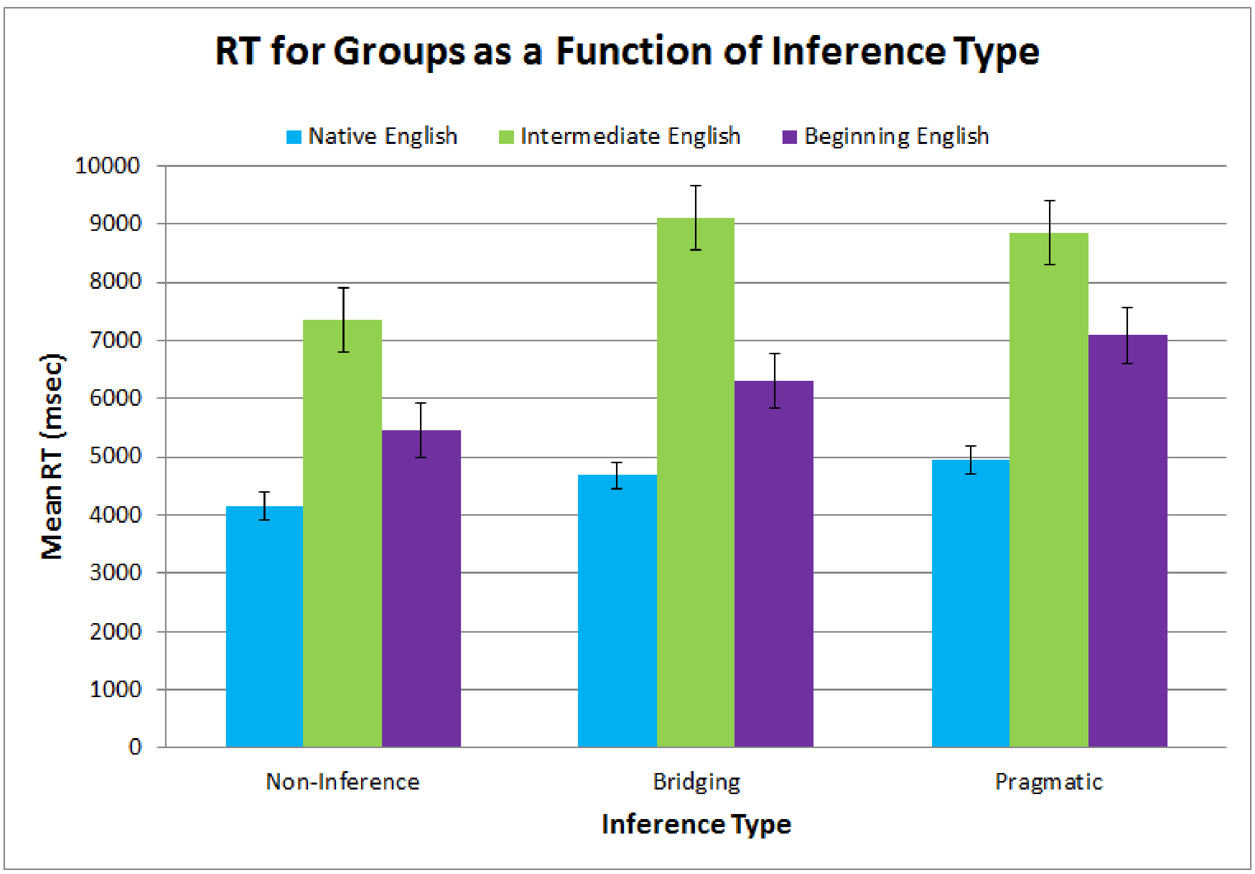

Figure 4. Mean Reaction Time for Groups as a Function of Inference Type

To further probe this interaction, one-way ANOVAs were conducted for group at each level of the Inference Type. Bonferroni tests revealed that for non-inferences, each group was significantly different from each other ($p<.05$), with the intermediate English readers actually taking the longest. This was consistent for the bridging inferences and pragmatic inferences as well, as each group was significant different from the other ($p<.05$), with the intermediate English readers again responding the slowest. This is also important in that it shows that although RT generally increases as the inferential complexity becomes greater, this is especially so for the pragmatic inference condition, where even the two lower proficiency groups become similar in the responding.

Additionally, there was a Group x Cognitive load interaction $F(2, 135) = 5.15$, $p= .007$, Cohen's $f = .10$, as seen in Figure 5. To further probe this interaction, one-way ANOVAs were conducted for group at high and low cognitive load. Bonferroni tests revealed that for the both the no load and the load condition, all three groups were significantly different from each other ($p<.05$), with the intermediate group responding a bit slower the other two groups. This is similar to the inference type x group interaction wherein the native English readers were different from the L2 groups under differing conditions of difficulty. This shows that difficulty level in general does seem to affect both effectiveness and efficiency.

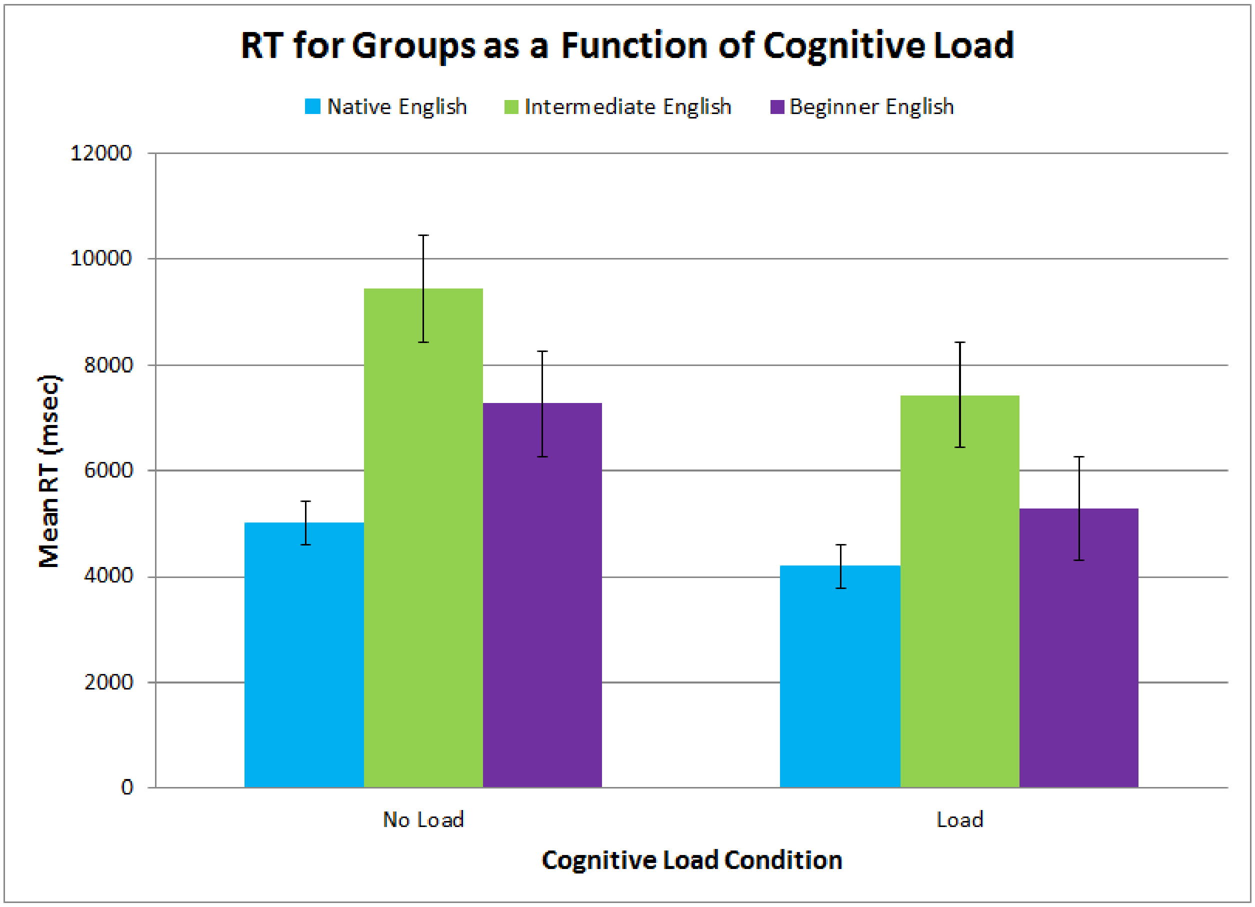

Figure 5. RT for Groups as a Function of Cognitive Load

It is quite interesting that under conditions of Cognitive Load all three groups responded faster than when under no cognitive load conditions. Perhaps the task in of itself is so taxing, that there is a spillover effect of the tones from the story to the questions, where participants may have felt the need to respond faster. This difference in the two load conditions was greatest for the intermediate English readers. Not surprisingly, accuracy did decrease as cognitive load increased (especially for the intermediate English readers). Maybe if participants had spent a bit more time in responding to questions in the cognitive load conditions, they would have performed better for those questions.

Finally, as displayed in Figure 6 there was a small but significant Inference Type x Cognitive Load interaction F (2, 270) = 4.35, p = .014, Cohen's f = .090.

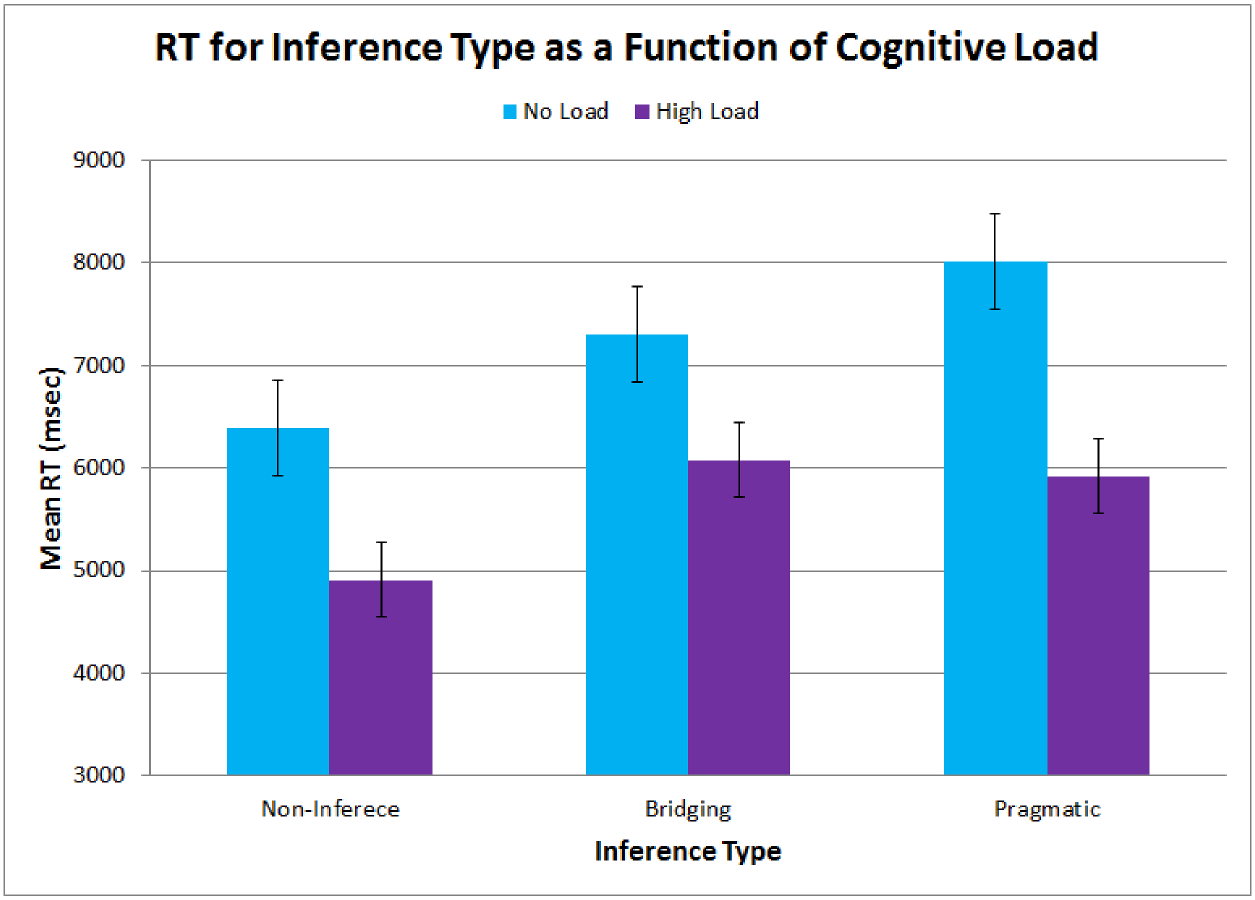

Figure 6. Reaction Time for Inference Type as a Function of Cognitive Load

Paired sample t-tests revealed that cognitive load differentially impacted the RT under non-inference $t(141) = 5.35$, $p < .001$, bridging inference $t(140) = 4.96$, $p < .001$ and pragmatic inference $t(138) = 6.81$, $p < .001$ conditions. It is once again interesting to note that RT decreased with cognitive load instead of increasing, as one might expect. The biggest drop was for the pragmatic inference condition, suggesting that when items are too difficult, perhaps participants defaulted to going faster, if they felt pressured to respond quickly to the tones they had heard while reading the story. It is also possible that since the cognitive load condition was in the last half of the experiment, perhaps participant fatigue was starting to be an issue where they again opted to go faster just to finish. However, overall, the pragmatic inferences still took longer, despite the fact that accuracy decreased for the most difficult conditions. Perhaps this is the reason that accuracy was decreasing, because the participants went too fast under these difficult conditions (although overall accuracy in general was still well above chance).

Chapter 5 - General Discussion

The purpose of the current experiments was to ask the questions of how the study of inference processing under different conditions of varying cognitive demands in L1 and L2 readers can contribute to our understanding of the episodic buffer. Reading in an L1 with or without adequate domain knowledge, and reading in English as an L2 at varying levels of proficiency both require the retrieval of information from LTM, which the episodic buffer then may bind to. As such, domain knowledge and reading proficiency levels were seen as proxies for difficulty level to understand how the episodic buffer plays a role in inferential reading.

For Experiment 1, although there were not many significant results, the ones that were significant, including some promising trends allows us to make some very tentative inferences on which of the three possible outcomes may be at work here. Overall, performance generally decreased in terms of effectiveness for all participants, suggesting that the central executive is certainly at work where performance systematically goes down for all (similar to outcome 3). However, given that of the three possible outcomes, although results seemed to be most consistent with the second outcome of the 'rich-get richer' hypotheses (Hambrick & Engle, 2002; Hambrick & Oswald, 2005; Meinz et al., 2011), suggesting that having adequate DK is good, but having sufficient WM resources just facilitates high domain knowledge even further. This may be why DK predicted variance as comprehension began to get more difficult, but topped out at the pragmatic inference no load, instead of pragmatic load condition.

Experiment 2 also seemed to walk the line between outcome 3, with the potential for outcome 2 of the rich-get-richer hypotheses. As the difficulty increases to the condition of greatest difficulty (pragmatic inference with cognitive load), the episodic buffer helps the native

English readers to maintain accurate performance. For the groups that are less proficient, (such as intermediate readers), especially when cognitive load is present, the episodic buffer may be needed even at the lower levels of inferential comprehension (fact and bridging inferences). For the beginner English readers, at the lowest proficiency levels in this study, it is not surprising that they would need to use all of their resources even for the easiest tasks, which for this group may have been quite difficult already, and as such they do not have sufficient working memory resources to be available for the more difficult levels. It is also important to note, that for both of the non-native English reading groups, working memory is at play when there is a cognitive load present. Interestingly, for reading comprehension efficiency (RT), the episodic buffer may be needed to bind information at even the earliest levels of processing, to allow participants to respond faster without compromising accuracy. The most proficient native English reader (rich getting richer) group is already at such a high level that they do not have a taxed central executive and thus do not exhibit effects of WM on RT, whereas the episodic buffer might be more important for the least proficient group at the easier stages of comprehension.

Overall, for both experiments, these results do not mean that difficulty level in itself does not impact comprehension in general. In fact, overall accuracy decreased as inferential complexity increased, and when cognitive load was present, typically with the native English readers being significantly better in performance than the two L2 groups (Experiment 2). This suggests that reading comprehension, even if one had adequate DK or native like proficiency, is still not immune from a taxed CE, decreasing overall effectiveness, although the episodic buffer can help compensate for some complexity in comprehension or difficulty of material.

Surprisingly, for both studies when cognitive load was present, RT actually decreased, which is opposite of results predicted by a speed-accuracy tradeoff. Intuition would tell us that

RT should increase as cognitive load is increased. This result may seem quite inconsistent with other RT studies, as well as with the predictions of the Attentional Control Theory (Calvo, et al., 1994; Darke, 1988b; Eysenck et al., 2007; Rai et al., 2011).

However, a possible explanation for this reduced RT in the high load condition may be the following: The Attentional Control Theory (as well as the constructivist approach in general) suggest that as difficulty increases and if there is not enough time for an individual to adequately process text, then higher level (i.e. pragmatic) inferences will not be drawn. The key is 'not enough time.' Perhaps the task itself was so taxing in the load condition that there was a spillover effect of the responding to the tones in the story to answering the subsequent questions, where participants may have felt the need to keep responding quickly (as they had to respond to the tone quickly). As such, they may have created their own 'forced speeded response' under the load condition even though they were not explicitly told to do so. Since accuracy did decrease as cognitive load increased for all participants in both studies, it is possible that the fast responding could have contributed to that.

Another possible explanation why participants were faster under conditions of cognitive load is related to the rich-get-richer hypothesis. Accuracy did decline under cognitive load, but was still quite high for all three reading groups. This suggests that perhaps the episodic buffer is able to adequately use top-down processing by retrieving relevant information from LTM, creating situation models that allow for speeded responses, without compromising accuracy. This still does not explain why RT was greater when no load was present, but perhaps in that condition they did feel like they could respond without being forced, and so the episodic buffer would not have to come into play, as at those levels, the CE can handle the demands in the no-load condition. Considering that RT does decrease in the load conditions, perhaps this is also evidence for the automatic processing by the episodic buffer in which responses can be made quickly, without severely compromising accuracy if there are enough cognitive resources to do so. Perhaps this is also evidence that under some circumstances the episodic buffer acts more like LTWM, working automatically, whereas in others it is uses more conscious processing (as for accuracy).

A third possible explanation might be a practice effect. Since the load condition was always the second half of the study, participants may have gotten faster in answering the questions with practice alone. However, accuracy did not seem to get better over time, as one would expect with the possible practice effect, indicating that this explanation, although possible, is less likely than those mentioned above.

Theoretically, these studies suggest that perhaps the episodic buffer can be a useful construct in understanding L2 reading (which according to the LTWM camp cannot be a part of their LTWM model, since reading is not an automatic task). The buffer itself maybe an addition to the WM model in line with several other models that have been discussed (i.e. LTWM, Cowan's working memory model), but it seems to have value in helping better understand L2 learning. It also suggests that beyond that, the EB can also be useful to account for the effect of DK also, as mentioned above, since both might be using the same underlying processes of retrieval of information from LTM to be processed by the episodic buffer and further manipulated by the CE.

Limitations, Applications, and Future Research

Although these studies have presented several important insights into the episodic buffer, there were several issues that may have limited the results. To begin with, future studies must utilize the linear mixed models approach in order to directly compare the results to the three possible outcomes instead of inferring their roles. Such statistical techniques may also help parse out the details of the underlying mechanisms of WM in reading. Additionally, for both studies, it may have been more beneficial to test groups that were truly more different from each other (e.g. specifically target baseball players, as experts very high in DK). It is also possible that testing the domain knowledge of baseball may be less useful than it has been in the past, as other sports are becoming more popular these days, such as basketball or football, which are typically the most well-liked and understood sports in NCAA schools. Future studies may then benefit from finding another domain in which the general population might have a greater variance of knowledge between high and low domain.

More specifically, as seen by several of the Experiment 2 results, the L2 groups typically differed from the native English readers, but much less so from each other. This suggests that although the two L2 groups were supposed to be at different proficiency levels, based on their class placement test and then LDT scores, perhaps they were more homogenous than originally anticipated. This may also explain why there was not any major impact of the LDT (a supplemental proficiency test in this study), because the differences between the LDT scores between the groups were small (Table 5); The groups were split with high and low LDT, instead of the initial groups as they were originally tested in, but they were still not very different from each other. Future studies may combine the two L2 groups together, or, as mentioned earlier, solicit groups that are more distinctly different from each other. It is also possible that since there was so much variability in each class tested (many countries of origin of students), this may have complicated the results. Alternative ANOVAs conducted on the groups split the two L2 groups into those using the Latin alphabet in their L1 (i.e. Spanish, Portuguese) (N=36) and those using other scripts (i.e. Chinese, Arabic, Japanese) (N=64), instead of by their ELP classes, suggested no differences between the language groups. Similar results were obtained when ANOVAs were conducted on the two L2 groups split into their ELP classes. Again, all of these suggest that a relative lack of diversity in proficiency between the two L2 proficiency groups may have accounted for some of the non-significant results.

Additionally, although significant results were obtained in this study, perhaps having stories that are different in their genre (e.g., expository prose instead of fairy tale-like stories) might produce different results, because in many cases in higher education, textbooks are more expository for information sake, rather than narrative for entertainment sake. This may make the text more familiar to readers, and elicit results more similar to those that students may experience in their actual learning.

Finally, this study mentioned LTWM when discussing the episodic buffer, although the former was never directly tested and the two were never directly compared. Future studies might search for an empirical distinction between LTWM and the episodic buffer. This would certainly help clarify the functioning of the episodic buffer, this new component of the WM model, but also perhaps allow the two theories (WM and LTWM) to be better competitively tested. For instance, if domain knowledge is tapped into by students learning a foreign language, cognitive load/dual-task paradigms may allow for better understanding of how these constructs work, especially with respect to second language learning. Additionally, this may also suggest that those who are teaching L2 learners, and perhaps even any higher education class, might have students practice WM strengthening skills, so that when they are under test situations, they can have their episodic buffers help overcome some of the load they may feel (for instance when dealing with timed tests, new topics, and so forth). In fact, several studies have suggested that

WM training can improve several cognitive skills such as reading (Chien & Morrison, 2010; Dahlin, 2011; Morrison & Chien, 2011), although much work needs to be done in this area.

Despite these limitations, these studies are also an important step in better understanding the episodic buffer, the newest addition to Baddeley's working memory model, for L1 readers with differing levels of DK, and especially in the area of L2 reading, where much empirical work still remains to be done. This is indeed a first step in addressing Juffs and Harrington's (2011) statement, "to date there has been no research on the possible role of the episodic buffer in L2 learning and use." Now there has been.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General,117*, 288–318.

Ackerman, P. L., & Kyllonen, P. C. (1991). Trainee characteristics. In J. E. Morrison (Ed.), *Training for performance: Principles of applied human learning* (pp. 193–229). Chichester, UK: Wiley.

Alderson, M.R., Kasper, L.J., Patros, C.H., Hudec, K.L., Tarle, S.J. and S. Lea. (2014). Working memory deficits in boys with attention deficit/hyperactivity disorder: An examination of orthographic coding and episodic buffer processes. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood Adolescence*, DOI: 10.1080/09297049.2014.917618.

Allen, R.J., Hitch, G.J. & Baddeley, A.D. (2009). Cross-modal binding and working memory. *Visual Cognition, 17*, 83-102

Alptekin, C., & Ercetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading,* 33(2), 206-219.

Ardila, A. (2003). Language representation and working memory with bilinguals. *Journal of Communication Disorders, 36,* 233-240.

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130(2)*, 224-237.

Baddeley, A.D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A.D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4(11),* 417-423.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews: Neuroscience, 4,* 929-939.

Baddeley, A.D. (2007). *Working memory, thought and action.* Oxford: Oxford University Press.

Baddeley, A.D. (2010). Working memory. *Current Biology 20,* 136-140.

Baddeley, A. D. (2012). Working memory: Theories, models and controversies. *Annual Review of Psychology, 63,* 1-29. doi: 10.1146/annurev-psych-120710-100422.

Baddeley, A.D., Allen, R.J. &, Vargha-Khadem, F. (2010). Is the hippocampus necessary for visual and verbal binding in working memory? *Neuropsychologia 48,* 1089–95.

Baddeley, A.D., Hitch, G.J., & R.J. Allen (2009). Working memory and binding in sentence recall. *Journal of Memory and Language, 61,* 438-456.

Baddeley, A. D. & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology, 129(1),* 126-145.

Baddeley, A.D., & Hitch, G. J. (1974). Working memory. In Bower, G. (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory,* Vol. 111. New York: Academic Press.

Baddeley, A.D., & Wilson, B. A. (2002). Prose recall and amnesia: Implications for the structure of working memory. *Neuropsychologia, 40,* 1737–1743.

Baars, B.J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences, 7(4),* 166-172.

Baumeister, R.F. (1984) Choking under pressure: self-consciousness and paradoxical effects on incentives on skillful performance. *Journal of Personality and Social Psychology, 46(3)* 610-620.

Beilock, S.L. & Carr, T.H. (2005). When high powered people fail: Working memory and 'choking under pressure' in math. *Psychological Science, 16(2)*, 101-105.

Berlingeri, M. et al., (2008). Anatomy of the episodic buffer: A voxel-based morphmetry study in patients with dementia. *Behavioral Neurology, 19,* 29-34.

Blake, M.L. (2009a). Inferencing processes after right hemisphere brain damage: Maintenance of inferences. *Journal of Speech, Language, and Hearing Research, 52*, 359-372.

Blake, M.L. (2009b). Inferencing processes after right hemisphere brain damage: Effects of contextual bias. *Journal of Speech, Language, and Hearing Research, 52*, 373-384

Britton, B. K., Stimson, M., Stennett, B., & Gülgöz, S. (1998). Learning from instructional text: Test of an individual-differences model. *Journal of Educational Psychology, 90,* 476–491.

Calvo, M. G. (2001). Working memory and inferences: Evidence from eye fixations during reading. *Memory*, 9(4-6), 365-381.

Cantor, J., & Engle, R. W. (1993). Working-memory capacity as long-term memory activation: An individual-differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1101–1114.

Chein, J.M., & Morrison, A.B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review, 17 (2),* 193-199.

Chiesi, H. L., Spilich, G.J., & Voss, J.F. (1979). Aquisition of domain-related information in relation to high and low domain knowlege. *Journal of Verbal Learning and Verbal Behavior, 18,* 257-273.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-185.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. Psychological Bulletin, 104, 163-191.

Dahlin, K.I. (2011). Effects of wokring memory training on reading in children with special needs. *Read Writ, 24,* 479-491.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

Daneman, M., & Merikle, P. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422-433. doi: 10.3758/bf03214546

Darke, S. (1988b). Effects of anxiety on inferential reasoning task performance. *Journal of Personality and Social Psychology, 55(3)*, 499-505.

De Pontes Nobre, A. et al., (2013). Tasks for the assessment of the episodic buffer: a systematic review. *Psychology and Neuroscience, (6) 3,* 331-343.

Dufour, R., & Kroll, J (1995). Matching words and concepts in L2 languages: A test of the concept mediation model of bilingual representation. *Memory and Cognition, 23 (2),* 166-180.

Ericsson, K.A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49*, 725–747.

Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211-245.

Estevez, A., & Calvo, M.G. (2000). Working memory capacity and time course of predictive inferences. *Memory, 8(1),* 51-61.

Eysenck, M.W., Santos, R., Derekshan, N., & Calvo, M.G.. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7(2),* 336-352.

Fincher-Kiefer, R., Post, T. A., Greene, T. R., & Voss, J. F. (1988). On the role of prior knowledge and task demands in the processing of text. *Journal of Memory and Language, 27,* 416-428.

Fitts, P., & Posner, M. I. (1967). *Human performance.* Belmont, CA: Brooks/Cole.

Franklin, S., Ramamurthy U., (2006) Motivations, Values and Emotions: 3 sides of the same coin. *Proceedings of the Sixth International Workshop on Epigenetic Robotics*, 41-48.

Graesser, A. C.. Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101,* 371-395.

Gray, R., Braver, T., & Raichle, M. (2002). Integration of emotion and cognition in the lateral prefrontal cortex. *Proceedings of the National Academy of Sciences, 99(6),* 4115-4120.

Guo, Y., & Roehrig, A.D. (2011). Roles of general versus second language (L2) knowledge in L2 reading comprehension. *Reading in a Foreign Language, 23(1),* 42-64.

Haenggi, D., & Perfetti, C. A. (1994). Processing components of college-level reading comprehension. *Discourse Processes, 17,* 83–104.

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology, 44,* 339–387.

Hambrick, D. Z., & Meinz, E.J. (2011). Limits on the predictive power of domain-specific experience and knowledge in skilled performance. *Current Directions in Psychological Science, 20(5),* 275-279. doi: 10.1177/0963721411422061.

Hambrick, D.Z., & Oswald, F.L. (2005). Does domain knowledge moderate involvement of working memory capacity in higher level cognition? A test of three models. *Journal of Memory and Language, 52*, 377–397.

Hammadou, J. (1991). Interrelationships among prior knowledge, inference, and language proficiency in foreign language reading. *The Modern Language Journal*, *75(1),* 27-38.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14,* 25-38.

Harrington, M., Ingram, J., & Proctor, M. (2014). *Multimedia assessment program languagemap.com*.

Harris, R. J., & Monaco, G. E. (1978). Psychology of pragmatic implication: Information processing between the lines. *Journal of Experimental Psychology: General,* 107, 1-22.

Henry, L. A. (2010). The episodic buffer in children with intellectual disabilities: An exploratory study. *Research in Developmental Disabilities, 31,* 1609-1614.

Hobbs, J. R. (1977). Pronoun resolution. *SIGART Newsletter, 61*, 28.

Horiba, Y, (1996). Comprehension processes in L2 reading: Language competence, textual coherence, and inferences. *Studies in Second Language Acquisition, 18(4),* 433-473.

Hulstijn, J.H. & Bossers, B. (1992) Individual differences in L2 proficiency as a function of L1 proficiency. *European Journal of Cognitive Psychology, 4(4),* 341-353.

Hunt, E. (1978). Mechanics of verbal ability. *Psychological Review*, 85, 109–130.

Huntley, J.D., & Howard, R.J. (2009). Working memory in early Alzheimer's disease: A neuropsychological review. *International Journal of Geriatric Psychiatry, 25*, 121-132. doi:10.1002/gps.2314.

Ikeda, M., Iwanga, M., Seiwa, H. (1996). Test anxiety and working memory system. *Perceptual and Motor Skills, 82*(3, Pt. 2), 1223-1231

Jeffries E., Lambon, R.A., Baddeley, A.D. (2004). Automatic and controlled processing in sentence recall: the role of long-term and working memory. *Journal of Memory and Language, 51,* 623–643.

Johnson-Laird, P.N. (1993). The mental representation of the meaning of words. In, A.I. Goldman (Ed.) *Readings in Philosophy and Cognitive Science.* (pp. 561-583). Cambridge, MA: MIT Press.

Juffs, A. (2004). Representation, processing, and working memory in a second language. *Transactions of the Philological Society*, *102(2),* 199-225.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching, 44,* 137-166.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122-149.

Karlsen, J.P., Allen, R.J., Baddeley, A.D., & Hitch, G.J. (2010). Binding across space and time in visual working memory. *Memory and Cognition, 38(3),* 292-303. doi:10.3758/MC.38.3.292

Kellogg, R. (2001). Long-term working memory in text production. *Memory & Cognition, 29*(1), 43-52.

Kembo, J.A. (2001). Testing of inferencing behaviour in a second language. *International Journal of Bilingual Education and Bilingualism, 4(2*), 77-96.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49, 294-303.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* New York: Cambridge University Press.

Kispal, A. (2008). *Effective teaching of inference skills for reading*. National Foundation for Educational Research. Research Report DCSF-RR031, 1-62.

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research, 18*, 141–175.

Leeser, M.J. (2007). Learner based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory. *Language Learning*, 57(2), 229-270.

Linderholm, T. (2002). Predictive inference generation as a function of working memory capacity and causal text constraints. *Discourse Processes, 34*(3), 259-280.

Linderholm, T. & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology, 94(4),* 778-784.

Luck, D., Danion, J., Marrer, C., Pham, B., Gounot, D., & Foucher, J. (2010). Abnormal medial temporal activity for bound information during working memory maintenance in patients with schizophrenia. *Hippocampus, 20,* 936-948. doi: 10.1002/hipo.20689

MacIntyre, P.D. & Gardner, R.C. (1994a). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, 44(2), 283-305.

MacIntyre, P.D., & Gardner, R. C. (1994b). The effects of induced anxiety on three stages of cognitive processing in second language learning. *Studies in Second Language Acquisition, 16(1),* 1-17.

MacIntyre, P. D., Noels, K. A., & Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning, 47*, 265–287.

Marmolejo-Ramos, F., et al. (2009). Reading between the lines: The activation of background knowledge during text comprehension. *Pragmatics & Cognition*, 17, 77-107.

McAnally, K., Castles, A. & Bannister, S. (2004). Auditory temporal pattern discrimination and reading ability. *Journal of Speech, Language, and Hearing Research, 47,* 1237-1243. doi:10.1044/1092-4388(2004/093)

McNamara, D. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38*(1), 1-30.

Meinz, E.J., Hambrick, D.Z., Hawkins, C.B., Gillings, A.K., Meyer, B.E., & Schneider, J.L. (2012). Domain knowledge and working memory capacity as predictors of components of skill in Texas Hold'Em poker. *Journal of Applied Research in Memory and Cognition, 1(1),* 34-40. DOI: 10.1016/j.jarmac.2011.11.001

Michael, E. B., & Gollan, T.H. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In Kroll, J.F., & DeGroot, A.M. (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*. (pp. 389-407). New York NY: Oxford University Press.

Miller, A., & Keenan, J. (2011). Understanding the centrality deficit: Insight from foreign language learners. *Memory & Cognition, 39*(5), 873-883. doi: 10.3758/s13421-010-0062z

Miyake, A., & Friedman, N.P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A.F. Healy, & L.E. Bourne (Eds.), *Foreign Language Learning: Psycholinguistic Studies on Training and Retention.* (pp. 339-364). Mahwah, NJ: Lawrence Erlbaum.

Miyake, A., Friedman, N., Emerson, M., Witzki A., and A. Howerter. (2000). The unity and diversity of executive functions and their contributions to "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100.

Morey, C.C. (2009) Integrated cross-domain object storage in working memory: Evidence from a verbal-spatial memory task. *The Quarterly Journal of Experimental Psychology, 62(11)*, 2235-2251.

Morrison, A.B. & Chein, J.M. (2011). Does working memory training work? The promise and challenges of enhancing cognition training by working memory. *Psychonomic Bulletin & Review, 18,* 46-60.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes, 26,*131–157.

Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English readings span tests. *Bulletin of the Psychonomic Society, 30,* 287-289.

Osaka, M., Osaka, N. & Groner, R. (1993). Language independent working memory: Evidence from German and French reading span tests. *Bulletin of the Psychonomic Society, 31,* 117-118.

Oulasvirta, A., & Saariluoma, P. (2006). Surviving task interruptions: Investigating the implications of long-term working memory theory. *International Journal of Human-Computer Studies, 64* (10), 941-961

Posner, M. I., & McLeod, P. (1982). Information processing models—In search of elementary operations. *Annual Review of Psychology, 33,* 477–514.

Perani, D. (2005). The neural basis of language talent in bilinguals. *Trends in Cognitive Sciences, 9(5)*, 211-213.

Rai, M. K., Loschky, L. C., Harris, R. J., Peck, N. R., & Cook, L. G. (2011). Effects of stress and working memory capacity on foreign language readers' inferential processing during comprehension. *Language Learning, 61*(1), 187-218. doi: 10.1111/j.1467-9922.2010.00592.x.

Rai, M. K. Loschky, L.C., & Harris, R.J. (in press). The effects of stress on reading: A comparison of first language versus intermediate second language reading comprehension. *Journal of Educational Psychology.*

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology, 80,* 16–20.

Ricks, T.R., & Wiley, J. (2009). The influence of domain knowledge on the functional capacity of working memory. *Journal of Memory and Language, 61,* 519-537. doi:10.1016/j.jml.2009.07.007

Rudner, M., & Ronnberg, J. (2008). The role of the episodic buffer in working memory for language processing. *Cognitive Processes, 9,* 19-28.

Service, E., Maury, S., & Luotoniemi, E. (2002). Individual differences in phonological learning and verbal STM span. *Memory and Cognition, 35*(5), 1122–1135.

Shimizu, M. (2005). Effects of a question on generating bridging inferences: L2 reading proficiency and a qualitative analysis of reading processes. *Bulletin of the Chubu English Language Education Society, 34*, 353-360.

Singer, M., Harkness D., & Stewart S. (1997). Constructing inferences in expository text comprehension. *Discourse Processes, 24,* 199-224.

Sluzenski, J., Newcombe, N. S., & Kovacs, S. (2006). Binding, relational memory and recall of naturalistic events: A developmental perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32* , 89-100.

Spilich, G. J., Vesonder, G. T., Chiesi, H. L., & Voss, J. F. (1979). Text processing of domain related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 22***,** 231–244.

Stahl, S.A., Hare, V.C., Sinatra, R., & Gregory, J.F. (1991). Defining the role of prior knowledge and vocabulary in reading comprehension: The retiring of number 41. *Journal of Reading Behavior, 23(4),* 487-508.

St. George, M., Mannes, S., & Hoffman, J. (1997). Individual difference in inference generation: An ERP analysis. *Journal of Cognitive Neuroscience, 9(6),* 776-787.

Unsworth, N., Heitz, R.P., Schrock, J.C., & Engle, R.W. (2005). An automated version of the operation span task. *Behavior Research and Methods, 37(3),* 498-505.

Van den Noort, M.W., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist, 11(4),* 289-296.

Vergauwe, E., & Cowan, N. (2014). A common short-term memory retrieval rate may describe many cognitive procedures. *Frontiers in Human Neuroscience*, 8 1-7. doi: 10.3389/fnhum.2014.00126

Vogel, E.K., Woodman, G.F., & Luck, S.J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance. 27,* 92–114.

Walker, C. H. (1987). Relative importance of domain knowledge and overall aptitude on acquisition of domain-related information. *Cognition and Instruction, 4,* 25–42.

Walter, C. (2004). Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory**.** *Applied Linguistics,* 25 (3), 315-339.

Was, C.A. (2014). An exploration into the evolution of working memory models. *International Review of Social Sciences and Humanities, 7(1),* 219-231.

Was, C.A., & Woltz, D.J. (2007). Reexamining the relationship between working memory and comprehension: The role of available long term memory. *Journal of Memory and Language, 56,* 86-102.

Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76,* 92-104.

Appendices

Appendix A: Demographic Questionnaire for Experiment 1

Age: _______ years
Place of birth (city, state/province,
country):____

Nationality (e.g., U.S., Mexican, Canadian): ____________________________________

Gender: M F

Year in school (circle one): Freshman Sophomore Junior Senior Graduate What is your major? __________________________________

A. Is English your **native language** (i.e., the very first language you learned)? Y N
If not, what is it? ______________________________

B. Please list all the languages you speak, understand, are formally studying, or have studied in the past:____________________________________

____

____

C. What is the highest level of English classes that you have taken? (i.e. English Comp I, etc).

____

D. If you are taking any foreign language classes on campus, which class, and what level are you currently in? _______________________________

E. How many hours per week do you currently spend in **a foreign language class**? _______hours per week

How many hours per week do you currently spend in the **language lab**? _______hours per week

How many hours per week do you currently spend doing **homework (outside of class/lab)?** ______ hours per week

circle one:

F. Not including your foreign language class/lab/homework, how often do you **hear** it spoken?

______hours per… day week month year

Not including your foreign language class/lab/homework, how often do you **speak** it? -

______hours per… day week month year

Not including your foreign language class/lab/homework, how often do you **read** it?

______hours per… day week month year

Not including your foreign language class/lab/homework, how often do you **write** in it?

______hours per… day week month year

G. Do you currently have any family or close friends who are fluent in a language other than English?

Y N If so, do you speak with them in that langauge? Y N

If so, how often? _____hours per… day week month year

circle one:

H. How would you rate your current **listening** skills in English?

Poor Fair Average Good Excellent

How would you rate your current **speaking** skills in English?

Poor Fair Average Good Excellent

How would you rate your current **reading** skills in English?

Poor Fair Average Good Excellent

How would you rate your current **writing** skills in English?

Poor Fair Average Good Excellent

I: If you know a language other than English (what language is it:________________________________)

How would you rate your current **listening** skills in that language?

Poor Fair Average Good Excellent

How would you rate your current **speaking** skills in that language?

Poor Fair Average Good Excellent

How would you rate your current **reading** skills in that language?

Poor Fair Average Good Excellent

How would you rate your current **writing** skills in that language?

Poor Fair Average Good Excellent

Appendix B: Questions on Baseball Knowledge Test (Spilich et, al. 1979)

1. If the visiting team is ahead, how many innings must a game go to be official?
2. What is the official score of a forfeit game?
3. If the batter makes an out in order to advance the runner one base, the play is termed a(n):
4. What does RBI stand for?
5. A pitcher warms up before going into the game. The place where he warms up is called the:
6. When a runner on base starts to run with the pitch and the batter knows he is going to run and tries definitely to hit the ball, this is called a(n):
7. Trying to get a runner form third base to home plate by bunting the ball is called a(n):
8. The distance between the pitcher's mound and home plate is:
9. The distance between any two bases is:
10. A pitcher who does not start the game but comes into a game at some point after it has started is called a(n) ________ pitcher.
11. How many innings must a starting pitcher pitch in order to receive credit for a win?
12. When a pitcher makes an illegal motion with a runner on base, the runner gets to advance a base. This advancing of a base due to a pitcher's illegal action is called a(n):
13. On each side of home plate there are rectangular markings where a batter stands when he is hitting. The marked off area is called the:
14. When a team is at bat, one person of that team stands in a rectangular area by first base and another person of that team stands in a rectangular area near third base. These areas are called:
15. With a man on first base the pitcher throws the ball and the catcher lets the ball get by him, allowing the runner to go from first to second base. The official scorer charges the ________ with a(n) ________.
16. This time the man on first base gets to second base because the pitcher threw the ball behind the batter and the catcher was unable to stop it. In this case the official scorer charges the ________ with a(n) ________.
17. An infielder goes "deep into the hole" and gets the batter out at first base. Which infielder did this?
18. If a ground ball hit by the batter rolls foul and then fair before the ball reaches third base, is the ball fair or foul?
19. When the pitcher suddenly turns and throws the ball to second base to try to make a runner out who is on second base, this action is called a(n):
20. If a batter is hit by a pitched ball, what do the rules say he is entitled to do?
21. What does ERA stand for?
22. A batter who is scheduled to bat immediately after the batter currently hitting is referred to as being:
23. In the American League, what does dh stand for?
24. A batter who hits either right- or left-handed is called a(n):

25. A batter hits a ground ball to the third baseman who in throwing the ball to first base throws the ball over the head of the first baseman and into the stands. Is the batter safe on first, safe on second, safe on third or does he get to score?
26. When the championship team of the American League pays the championship team of the National League it is called the World Series. What is the number of winning games for one team it takes to win the World Series?
27. A player who has not been in a game is put up to bat to hit for someone who has been in the game. The player put into the game is called a(n):
28. In baseball, one person determines whether a pitch is a ball or a strike. This person is a(n):
29. When is a batter allowed to run to first base even though he struck out?
30. With runners on second and third base, a batter may be walked intentionally. Assume the batter is not an outstanding hitter. What would be the major purpose of such a move?
31. What type of pitcher generally has the best "move" to first base to pick off runners?
32. Pitchers tend to throw "off-speed" or "change of pace" pitches when facing a ________ hitter.
33. If a strong left-handed batter is coming up to bat in a critical situation, is the manager of the team in the field more likely to bring in a left-hander or right-hander if he brings in a new pitcher?
34. If a third-baseman who is not noted for his speed is playing deep at third base, a batter may try to get to first base by:
35. One way in which a batter gains better control of his bat's trajectory (though sacrificing power) is to:
36. When trying to catch a hard-hit ground ball, what will a good short stop do to prevent the ball from getting into the outfield in the event that he can't catch it?
37. When chasing a deep fly ball, how can the outfielder tell he is getting close to the wall (so as not to crash into it) without taking his eye off the ball?
38. The most common method that pitchers use to make a ball curve is to ________ their ________ when releasing the ball.
39. A good base-stealer will make his "move" or "break" just as the pitcher:
40. When charging a well-placed bunt, the third-baseman will reduce the time it usually takes him to make the play by:
41. Without moving his feet, the best way for a right-handed batter to make sure he hits to the right side of the diamond is to:
42. When a ground ball looks like it may make some time consuming bounces, the infielder should ________ the ball so that a play can be made before the batter reaches first base.

Added Questions

46. Approximately how many baseball games do you attend each season?
47. During baseball season, approximately how many games do you watch on television each week?
48. On a scale from 1 to 10 (with 1 meaning "Not at all interested" and 10 meaning "Extremely interested"), how interested are you in baseball?
49. Did you ever play or coach for a community baseball or softball team?
50. Did you ever play or coach for a school baseball or softball team?

Appendix C: RCT materials for Experiment1 (Baseball stories)

Note that the first passage is an example of what participants saw in the actual RCT task. The story and each question appeared individually.
* Participants saw question with accompanying instructions. For the Appendix they are labeled Fact, Bridging Inference and Pragmatic Inference for the readers of this manuscript.
**Each question was followed by the same rating system as the first question. The end anchors were a different description pertaining to each specific question. The rating instructions will be available for each question on the screen.
***The correct answers are underlined only for the readers of this document. Participants only saw the likert scale with anchors in plain text.

Passage 1 (349 words)

Good evening baseball fans. Well, the temperature here at Senators stadium is a cool 60 degrees, and it's a perfect September night for baseball. The Senators and Redbirds are tied up at three runs apiece. The Senators are up to bat, and keep in mind that the winner of the game tonight will move into first place in the Western Division. Punch Grubb is pitching for the Redbirds and came into the game with a 2.45 earned run average and he has 19 wins on the year. Now Juan Sanchez steps up to bat. This speedy lead-off man has a .240 batting average and leads the team in stolen bases. While he steps away to check the grip on his bat, here's some news from around the league: the Cougars are leading the Robins by two in the bottom of the sixth. Now Grubb takes the sign. Here's the pitch, and a swing of the bat sends a line drive into centerfield. The leadoff man will hold up at first base with a single, and now we can see that Dave Madden is warming up in the bullpen for the Redbirds. This kid from the Bronx has been unbelievable this year with 38 saves and a 1.95 earned run average. Chip Craik is the next batter of the inning. This veteran is a scrappy clutch hitter with a .285 batting average, and he leads the team in walks. Now Grubb takes the sign and is getting ready to deliver—and around the infield, the third baseman creeps in toward home plate, and the shortstop shifts over to cover second base. Here comes the pitch, and there goes a bunt down the first baseline. Grubb charges in and fields the ball, and his only option is to put the batter out at first base. The crowd tonight just saw a textbook bunt, and the runner advances to the next base with no problem. Some people say that bunting is a lost art, but that was a good one. He just squared around and steered the ball right down the baseline.

Fact:*
Please provide a rating of how certain you are about the answer for each question.**
1 = strongly certain that it is the Redbirds
2= Mostly certain it is the Redbirds
3= somewhat certain it is the Redbirds
4= I Don't know
5= somewhat certain it is the Cougars
6= mostly certain it is the Cougars
7= Strongly certain that it is Cougars

The Senators are playing against which team?

<u>Redbirds ***</u> Cougars

1 2 3 4 5 6 7

Strongly certain it is the		Strongly certain it is the

If the Senators win, they will be atop which division?

Eastern <u>Western</u>

1 2 3 4 5 6 7

Which inning is the other game in?

<u>Sixth</u> Seventh

1 2 3 4 5 6 7

Bridging Inference

Please provide a rating of how certain you are about the answer for each question.*
1 = strongly certain that it is Juan Sanchez
2= Mostly certain it is the Juan Sanchez
3= somewhat certain it is the Juan Sanchez
4= I Don't know
5= somewhat certain it is the Dave Madden
6= mostly certain it is the Dave Madden
7= Strongly certain that is Dave Madden

Who went to check the grip on his bat?

<u>Juan Sanchez</u> Dave Madden

1	2	3	4	5	6	7
Strongly certain it is						Strongly certain it is

Who is from the Bronx?

Chip Craik						*Dave Madden*
1	2	3	4	5	6	7

Who has 19 wins on the year?

Dave Madden						*Punch Grubb*
1	2	3	4	5	6	7

Pragmatic Inference

Please provide a rating of how certain you are about the answer for each question

1 = strongly certain that it is NOT raining in the stadium
2= Mostly certain it is the NOT raining in the stadium
3= somewhat certain it is NOT raining in the stadium
4= I Don't know
5= somewhat certain it is raining in the stadium
6= mostly certain it is raining in the stadium
7= strongly certain that is raining in the stadium

It is raining at the stadium

1	2	3	4	5	6	7
Strongly						Strongly

Many players bunt as much as they used to

1	2	3	4	5	6	7

Pitchers like to pitch to Chip Craik

1	2	3	4	5	6	7

Rest of the stories continued. The responses were all on a 7 point scales as described above.

<u>Passage 2</u> <u>(333 words)</u>

It's the <u>bottom of the 7th inning</u>, on this beautiful <u>July day</u>. The Hurricanes are up to bat, and Skip Lawson is pitching for the Bluebirds this afternoon. <u>He gave up a home run in the fourth inning but has been unstoppable ever since with 7 strikeouts</u>. <u>Now Ray Nickerson steps up to bat. He's a left-hander pull hitter with a .325 batting average and 15 home runs on the year.</u> The rightfielder moves back towards the warning track, and the shortstop swings around to the right. One of the keys to <u>the Bluebirds' success in the first half of the season is solid play in the outfield. The Bluebirds' outfielders have only four errors on the year</u>. Lawson is ready now. Here comes a fastball, and a swing of the bat sends a looper into shallow leftfield, and that was only the third hit that Lawson has allowed all afternoon. <u>Gabriel Garcia</u>, the number seven hitter in the lineup, is next to bat. <u>He was out earlier this year with a pulled hamstring</u>, but has played solid baseball since. Now Lawson delivers—and there goes the runner, and a groundball is hit into rightfield. That was perfect execution. The batter holds up at first base with a single. <u>Sam Philipe is the next batter of the inning. This kid has struggled all season long with a .200 batting average and no hits for his last 15 times at bat</u>. Now he glances over at Buddy Brockman, the third base coach. <u>He had some health problems back in '95</u>, but he says he's never felt better. Here's the pitch from Lawson, and a weak pop-up is hit over to shallow rightfield. The first baseman backs up and makes the catch for the out, and the runners cannot advance. Well, this just in: the official attendance for today's <u>game is just under 30,000, and 5,000 of the fans today are summer school kids from around the metro</u>. We'll be back after this short break.

<u>Fact (non-inference)</u>

1. Which part of the 7th inning is it? a) Top *<u>b) Bottom</u>*
2. That game had how many total fans? *<u>a)30,000</u>* b)5,000
3. Which month is the game being played in? a) June *<u>b) July</u>*

<u>Bridging Inference</u>

1. Who was out with a hamstring injury? *<u>a)Gabriel Garcia</u>* b)Skip Lawson
2. Who is the left-handed hitter with 15 homeruns? a)Skip Lawson b)<u>Ray Nickerson</u>
3. Who had health problems in 1995? *<u>a)Coach Brockman</u>* b)Sam Philipe

<u>Pragmatic Inference</u>

1. The Bluebirds are a good defensive team
    a. strongly certain that bluebirds are NOT a good defense team
    b. <u>strongly certain that bluebirds are a good defense team</u>
2. Lawson is a bad pitcher?
    a. <u>Strongly Certain Lawson is NOT a bad pitcher</u>
    b. Strongly Certain Lawson is a bad pitcher
3. Sam Philipe might be traded soon?
    a. Strongly Certain that Sam Philipe will NOT be traded soon
    b. <u>Strongly Certain that Sam Philipe will be traded soon</u>

Passage 3 (348 words)

Hello again baseball fans. You're tuned into *WBLG*, the big-league radio network, and it's the bottom-half of the 1st inning. The Generals are up to bat, and if you just joined us, the Clippers jumped out to a 2-to-nothing lead in the top-half of the 1st inning. Remember that the winner of this game will take sole possession of second place in the Northern Division. Martin Sliwinski is pitching for the Clippers this afternoon, and he comes into today's game with two wins and two losses on the season and a 4.50 earned run average. Now Tom Wilcox steps up to bat. This veteran lead-off hitter has a .275 batting average, and once again he leads the league in stolen bases. Here comes the pitch from Sliwinski, and a blooper is hit into leftfield. Well Sliwinski threw a perfect pitch—a fastball high and inside—but sometimes the best pitch in the world isn't enough against a good batter. The fans are still pouring into this stadium, and it looks like we have an exciting game in store. Although it's still early in the season, both of these teams would like to get into position to make the playoffs. Tony Zonderman is the next batter of the inning. He has a .225 batting average, and comes into the game with no hits for his last 15 times at bat. You can bet that he would like to break out of this slump today. Now Sliwinski is ready, and he glances over at the runner. Here comes the pitch, and a swing of the bat sends a line-drive to the gap in left-centerfield. The leftfielder charges over to field the ball, and here comes the throw from the outfield. The runner slides in under the tag at home plate, and he scores as Sliwinski throws his glove to the mound. That throw was just a fraction too late, and the batter holds up at second base with a double. Now the 18,000 fans are on their feet, and what a beautiful day in the month of May for baseball.

Fact (non-inference)

1. Which division do the teams play on? *a) Northern* b) Western
2. Who scores first in this game? a)Generals *b) Clippers*
3. What radio station is airing the game? *a)WBLG* b)WMLB

Bridging Inference

1. How as 2 wins this season? *a) Martin Sliwinski* b) Tom Wilcox
2. Who leads the team in stolen bases *a) Tom Wilcox* b) Martin Sliwinski
3. Who has not had a hit in 15 games? a) Tom Wilcox b) Tony Zonderman

Pragmatic Inference

1. There two teams are fun to watch
   a. Strongly certain the two teams are NOT fun to watch
   b. Strongly certain the two teams are fun to watch
2. The clippers are the home team
   a. Strongly certain the Clippers are NOT the home team
   b. Strongly certain the Clippers are the home team
3. The pitcher was happy with the last pitch
   a. Strongly certain the pitcher was NOT happy with the last pitch
   b. Strongly certain the pitcher was happy with the last pitch

<u>Passage 4</u> <u>(306 words)</u>

There's not a cloud in the sky, and the temperature is hovering around 75 degrees. Now the young Keith Stanovich steps to bat. <u>This kid from Oklahoma</u> has demonstrated good power this year with a batting average of .305 and already 10 home runs. Here's the pitch from Perez, and a hard groundball is hit to the left side of the infield. <u>The shortstop fields the ball and makes the throw over to first base to put the batter out</u>, and the runner does not advance. While we have a second, here's some news from around the league: the Bearcats are leading the Blizzards 2–nothing in the bottom of the second inning, and the <u>Dukes</u> are trailing the Seminoles 4–1 early in the third. Back to the action as <u>Angel Cabrera</u> steps up to bat. This big right-hander has been unstoppable this year. He has a .350 batting average and <u>leads the team in RBIs</u>. The catcher steps out to the pitcher's mound for a chat with Perez, and what he's telling Perez is to keep his pitches away from the outside part of the plate. <u>The outfielders shift back and around to the left, and the shortstop moves into position for a pick-off</u>. Now Perez is ready. Here comes the pitch, and there goes a hit over the shortstop's head and into <u>centerfield</u>. Well that was obviously not what Perez wanted to do. The runner gets the signal to try for the score. Here comes the throw to home plate from the outfield, and he's safe. The batter holds up at second base with a standup double, and now the <u>pitching coach is clearly fuming as he is heading out to the mound for a chat with Perez</u> We'll take a short pause for station identification here on WBLG, the big-league radio network.

<u>Fact (non-inference)</u>

1. The first batter was safe on first base — *<u>a) Yes</u>* — b) No
2. Which other team is playing tonight? — *<u>a) Dukes</u>* — b) Angels
3. Where did the last batter hit the ball? — a) Left-field — *<u>b) Centerfield</u>*

<u>Bridging Inference</u>

1. Who is from Oklahoma? — *<u>a) Keith Stanovich</u>* — b) Angel Cabrera
2. Who leads the team in RBI's? — *<u>a) Angel Cabrera</u>* — b) Perez
3. Who came out to speak with Perez? — a)The manager — <u>b) The pitching coach</u>

<u>Pragmatic Inferences</u>

1. Stanovich is a good player
   a. Strongly certain Stanovich is NOT a good player
   b. <u>Strongly certain Stanovich is a good player</u>
2. The defense of afraid of Cabrera
   a. Strongly certain the defense is NOT afraid of Cabrera
   b. <u>Strongly certain the defense is afraid of Cabrera</u>
3. The pitcher may be substituted for a relief pitcher
   a. <u>Strongly certain the pitcher may NOT be substituted for a relief pitcher</u>
   b. Strongly certain the pitcher may be substituted for a relief pitcher

Practice passage (301 words)

The Ridgeville Robins are playing the Center City Cougars. The Robins are leading 5-3 with the Cougars at bat in the first half of the fifth inning. The sky is getting darker, and the rain is becoming heavier, typical of October. The Cougar's first batter, Harvey Jones is taking his time coming to the plate The umpire steps back from behind the plate and tells him to step into the batter's box. Jones, the hitter, is left handed and has a batter average of .310. Claresen, the pitcher has allows only four hits, has walked one, and has struck out six. This performance is about average for Claresen since this left hander has an earned-run average of 6.00 and typically strikes out quite a few batters. Claresen now adjusts his cap, touches his knee, begins his windup and delivers a high fastball that the umpire calls "Ball One." The Robins catcher, Don Postman returns the ball, and Claresen takes the sign. The next pitch is swung on and hit to centerfield. Maloney comes in and he catches it for the first out. The next batter is a powerful hitter, Fred Johnson who leads the league in home runs with 23. Claresen no doubt is glad to face him with no one on base. Claresen is now getting his sign from the catcher, begins his windup and throws a curveball breaking into a right-hand batter at the knees. The umpire calls it a strike. Calresen is now getting ready again, winds up and throws, and Johnson hits it off to the right and into the stands, a foul ball. Claresen rubs up the new ball, takes his sign and throws a fastball which just misses the bill of Johnson's cap. Johnson took one step toward the mound, but then came back. What a game!

Fact (non-inference)

1. What is the name of the Ridgeville team? *a) Robins* b) Cougars
2. What was the first pitch a) A ball b) A Strike

Bridging Inference

1. Who went back to the batter's box? A) the umpire b) Harvey Jones

2. Who made the first out? A) Postman b)Maloney

Pragmatic Inference

1. Johnson likes the pitcher
   a) Strongly certain Johnson does NOT like the pitcher
   b) Strongly certain Johnson does like the pitcher.

2. This game is a playoff game
   a) Strongly certain this is NOT a playoff game
   b) Strongly certain this is a playoff game

Appendix D: Demographic Questionnaire for Experiment 2

Age: _______ years

Place of birth (city, state/province, country):____

Nationality (e.g., U.S., Mexican, Canadian): ____________________________________

Year in school (circle one): Freshman Sophomore Junior Senior Graduate What is your major? ___________________________________

A. Is English your **native language** (i.e., the very first language you learned)? Y N If not, what is it? _______________________________

B. Please list all the languages you speak, understand, are formally studying, or have studied in the past:_____________________________________

____

____

C. What is the highest level of English classes that you have taken? (i.e. English readingI, etc).

____

D. If you are taking any foreign language classes on campus, which class, and what level are you currently in? ________________________________

E. How many hours per week do you currently spend in **a foreign language class**? _______hours per week

How many hours per week do you currently spend in the **language lab**? _______hours per week

How many hours per week do you currently spend doing English **homework (outside of class/lab)?** ______ hours per week circle one:

F. Not including your foreign language class/lab/homework, how often do you **hear** it spoken?

______hours per… day week month year

Not including your foreign language class/lab/homework, how often do you **speak** it? -

______hours per… day week month year

Not including your foreign language class/lab/homework, how often do you **read** it?

______hours per… day week month year

Not including your foreign language class/lab/homework, how often do you **write** in it?

______hours per… day week month year

G. Do you currently have any family or close friends who are fluent in a language other than English?

Y N circle one:

If so, do you speak with them in that langauge? Y N If so, how often? _____hours per…

day week month year

circle one:

H. How would you rate your current **listening** skills in English?

Poor Fair Average Good Excellent

How would you rate your current **speaking** skills in English?

Poor Fair Average Good Excellent

How would you rate your current **reading** skills in English?

Poor Fair Average Good Excellent

How would you rate your current **writing** skills in English?

Poor Fair Average Good Excellent

I: If you know a language other than English,

How would you rate your current **listening** skills in that language?

Poor Fair Average Good Excellent

How would you rate your current **speaking** skills in that language?

Poor Fair Average Good Excellent

How would you rate your current **reading** skills in that language?

Poor Fair Average Good Excellent

How would you rate your current **writing** skills in that language?

Poor   Fair   Average   Good   Excellent

J. What was your TOEFL score? ______________

K. Are you transferring to K-State from another university? Y N   If yes, which one?

____

L. Which is the highest English reading class you have taken? ________________________

Appendix E: RCT materials for Experiment 2 (short stories) from Rai et al., 2011

Note: Format was the same as in Appendix C

There are fifty students in young Billy's economics class. Unfortunately, the four most annoying people in the class happen to sit nearby. The student who sits to his right, John, props his book up on his desk and sleeps through the class. Sam, who sits to Billy's left, fidgets a lot. He spends the class period practicing tricks he can do with his pen. About every half minute, he drops it and has to look for it. Mike, who sits in front of Billy, reads the newspaper all through class. There just can't be that much interesting news to fill the whole hour! And the red-haired kid who sits behind him constantly kicks his chair. It's as if the red-haired kid and his friends are trying to make Billy fail his class! Every time he kicks Billy's chair, Billy remembers how much he can't stand economics.

Non-inference/fact

Please answer how certain you are of what Billy is studying by pressing the number based on the rating scale below.

1=strongly certain Billy is studying Current events
2-=mostly certain Billy is studying Current events
3= somewhat certain Billy is studying Current events
4= I don't know
5= somewhat certain Billy is studying Economics
6= mostly certain Billy is studying Economics
7= strongly certain Billy is studying Economics.

What is Billy studying?

Current events Economics

1 2 3 4 5 6 7

Strongly certain Billy is studying Current

Strongly certain Billy is studying

Bridging Inference:

Who kicks Billy's chair?

the red-haired kid — Mike

1	2	3	4	5	6	7
Strongly certain the red-haired kid kicked Billy's						Strongly certain Mike kicked Billy's

Pragmatic Inference:

Does Billy like the people sitting near him?

1	2	3	4	5	6	7
Strongly certain Billy does NOT like people sitting near him						Strongly certain Billy does like people

Rest of the questions for passage above.

Fact (Non-Inference):
How many students were in the class a) 30 b) 50
What does Mark read through class? a) Newspaper b) Comic books

Bridging Inference
Who props books on his desk? a) Rick b) Billy
Who practiced tricks with his pen during class? a) Mark b) Billy
Pragmatic Inference
Did the teacher criticize the bad behavior of the boys?
a) Strongly certain that the teacher did NOT criticize the bad behavior of the boys
b) Strongly certain that the teacher did criticize the bad behavior of the boys

Is Sam attending a University?
a) Strongly certain that Sam is NOT attending a university
b) Strongly certain that Same is attending a university

Passage 2

Once upon a time there was a very beautiful doll's house made of red bricks and with white windows. It had real curtains and even a chimney. It belonged to two dolls named Sally and Maria. Maria was the cook and always cooked delicious food. For dinner, there were two red lobsters, a ham, a fish, a pudding, and some pears. They would not come off the plates, and they were extremely beautiful. One morning the two dolls had gone out for a drive. There was a little scratching noise in a corner near the fireplace, where there was a hole. Out came two mice, Tom and Anna, Toms' wife. Tom and Anna went upstairs, peeped into the dining room and squeaked with joy at the sight they saw. Meanwhile, Sally and Maria were driving toward home. Then, they arrived back at the house. As they went inside, they both screamed when they saw who was in their house. Anna screamed too as she ran away before getting in trouble.

Non-Inference (Fact)

- Who was Anna? a. another doll *b. Tom' wife.
- Where was there a hole? *a. near the fireplace b. in the kitchen
- What color were the bricks on the house? a. white *b. red

Bridging

- Who arrived back at the house? a. Tom and Anna *b. Sally and Maria
- What belonged to Sally and Maria? *a. the house b. the mice
- Who ran away when the owner's came home? *a) Anna b) Tom

Pragmatic Inference

Did Tom and Anna live downstairs

a) Strongly certain that Tom and Anna did NOT live downstairs
b) *Strongly certain that Tom and Anna did live downstairs

Was the food on the plates real?

a) *Strongly certain that the food on the plates was NOT real
b) Strongly certain that the food on the plates was NOT real

Were Sally and Maria expecting to find someone when they came home?

a) *Strongly certain that Sally and Maria were NOT expecting to find someone when they came home.
b) Strongly certain that Sally and Maria were NOT expecting to find someone when they came home.

Passage 3

Kate and Mike were firefighter friends. Every sunny summer morning, Miguel would climb to the top of the fence between their yards and shout, “Fire! Fire!” Then Kate would jump and yell, “Heroines and heroes to the rescue!” They would ride into danger and save many people’s lives. But one day, Kate said, “I want to do something different, like look for golden treasure buried under the sea… or fly away in this magic tree.” “Wow,” said Miguel, “we’ve never played that kind of game before.” “Then climb on up,” said Kate, “because this tree is ready to ROAR!” “Help!” cried Miguel. “He’s an evil dragon of the sea, and I am in trouble!” Kate grabbed the tail and climbed up on the slippery scales. Then she tickled him until he set her friend free. All of a sudden thunder sounded in the sky above and it began to rain. “Oh no!” yelled Mike as he ran inside crying that he couldn’t play outside anymore.

Fact:

What did Kate want to look for under the sea? a. a dragon *b. golden treasure
What did Mike climb every summer morning? *a. the fence b. the tree
What were Mike and Kate pretending to be at the beginning of the story?
*A. Firefighters b. Police officers

Bridging:

Whom did Catalina tickle? a. Mike *b. the dragon
Who rode into danger and saved lives? *a Mike and Kate b. the magic dragon
Who ran inside when it started to rain? *a. Mike b. Kate

Pragmatic:

Were Kate and Mike neighbors?

a) Strongly certain Kate and Mike were NOT neighbors
b) *Strongly certain Kate and Mike were neighbors

Were Kate and Mike adults?

a) *Strongly certain Kate and Mike were NOT adults
b) Strongly certain Kate and Mike were adults

Did Mike and Kate like to play inside?

a) <u>Strongly certain Mike and Kate did NOT like to play inside</u>
b) Strongly certain Mike and Kate did like to play inside

Passage 4

One day a man was walking through a forest and heard a bird crying. The bird then said, "I am a king's daughter by birth, and am bewitched, but you can set me free." "What am I to do?" he asked. She said, "Go further into the forest, and you will find a house, wherein sits an old woman, who will offer you food and drink, but you must accept nothing, for if you eat and drink anything, you will fall into a sleep. I will come in a carriage for three days straight every afternoon at two o'clock. The carriage will have three white horses. Each of the three times when I arrive, you must kiss me on the cheek, and then I will disappear." He promised the bird to do everything that she desired, although he also remembered the old woman. But the bird said, "Alas, I know already that you will not set me free, you will accept a drink from her."

Fact:

How many horses will the carriage have? a. four *b. three

What time did the bird promise to come every day? *a. two o'clock b. three o'clock

Where does this story take place? *a. In the forest b. at a castle

Bridging:

From whom will the man accept a drink? *a. the old woman b. the king's daughter

Who was a king's daughter by birth? a. the old woman *b. the bird

The old man promised to do everything who desired? *a) the bird b. the old woman

Pragmatic:

Did the man want to help the bird?

a) Strongly certain the man did NOT want to help the bird
b) *Strongly certain the man did want to help the bird

Had the old woman cast a spell on the king's daughter?

a) Strongly certain the old woman had NOT cast a spell on the king's daughter
b) *Strongly certain the old woman had cast a spell on the king's daughter

Did the bird trust the old man?

a) *Strongly certain the bird did NOT trust the old man
b) Strongly certain the bird did trust the old man

Passage 5

Every Wednesday night I like to watch my favorite show about the silly Doctor Henry_Juraska. I usually get together with my two best friends, Steve and Nancy, and we like to make fun of the TV show. Often we get carried away and start laughing so hard that we miss some of the lines. It's great because it gives us a chance to relax and forget about our worries from that week. It reminds me of the saying that laughter is the best medicine. For example, one time I had a terrible headache but after my weekly laughing session with Dr. Juraska I felt better. I always feel better after my weekly appointment with him.

Fact:

The show was on every: *a. Wednesday night b. Thursday night
What type of illness did the narrator have one time? *a. headache b. stomach ache
What is the best medicine? A. relaxing *b. laughter

Bridging:

The author felt better after his weekly appointment with a.Steve *b. Dr. Juraska
Who could relax and forget about worries? *a. Steve and Nancy b. Dr. Juraska
Who gets carried away that they miss lines? *a. The author and his friends b. Dr. Jurska

Pragmatic:

The show the author was watching is a comedy.

a) Strongly certain the author was NOT watching a comedy
b) *Strongly certain the author was watching a comedy

Was Dr. Juraska a real doctor?

a) *Strongly certain Dr. Juraska was NOT a real doctor
b) Strongly certain Dr. Juraska was a real doctor

The author has a stressful life?

a) Strongly certain the author does NOT have a stressful life
b) *Strongly certain the author does have a stressful life

Passage 6

Yesterday when I went to the grocery store, I saw some friends that I haven't seen since high school. As I was looking at the many ice cream flavors, I felt someone hug me from behind, and it was Carol, a good friend of mine whom I met in grade school. Valerie and Linda were following close behind. Valerie and I met in choir, despite the fact that neither of us had much music talent. The way I met Linda is sort of funny- I ran into her one day in the school cafeteria while I was carrying my lunch tray. Valerie and Carol then showed up and quickly defused the situation. From that moment on, the four of us became best friends all through high school. It was great seeing Carol again. The memories of how we all met rushed through my mind as she hugged me. After meeting at the grocery store, Carol said we should all have dinner tonight. She will pick me up in her car at 8pm. I'm very excited to see the girls again.

Fact:

Where was the narrator hugged? a. the school cafeteria *b. the grocery store
Where did Valeria and the narrator meet? a. in the cafeteria *b. in choir
What was the narrator doing when he was hugged? *a. looking at ice cream flavors b. paying for his items

Bridging:

Who hugged the narrator? *a. Carol b. Linda
Whom did the narrator meet by running into her in the cafeteria? a. Valerie *b. Linda
Whose car will the narrator be picked up in? *a. Carol b. Linda

Pragmatic:

Did the narrator spill her food at school?

a) Strongly certain the narrator did NOT spill her food at school
b) *Strongly certain the narrator did spill her food at school

Had it been many years since the characters were in high school?

a) Strongly certain it had NOT been many years since the characters were in high school
b) *Strongly certain it had been many years since the characters were in high school

The narrator likes the people he met at the grocery store.

a) Strongly certain the narrator does NOT like the people he met at the grocery store
b) *Strongly certain the narrator does like the people he met at the grocery store

Practice

In a village lived an old shoemaker. He made a pair of little shoes out of some old pieces of red cloth. They were intended for a little girl, whose name was Cristina. Cristina received the shoes and wore them for the first time on the day of the village elder's funeral. Although they were not suitable for mourning, nothing else was available. Partway along the walk to the church, a large old carriage came by, carrying an old lady, a clergyman and a tiny baby girl. The old lady looked on the scene and took pity. She said to the clergyman, "We should show charity and share our carriage with those in need." And so the old lady opened the door of her carriage for Cristina. Upon arriving at the funeral, the old lady was very kind. However, she said the red shoes were inappropriate and gave her a new pair.

Fact:
What was the sex of the tiny baby? *a. female b. male
What color were Cristina's shoes? *a. red b. black

Bridging:
Who received a new pair of shoes? *a. Cristina b. the old lady
Who suggested to the clergyman that they should show charity? a. Cristina *b. the old lady

Pragmatic:
Did Cristina walk all the way to the funeral?
a) *Strongly certain that Cristina did NOT walk all the way to the funeral
b) Strongly certain that Cristina did walk all the way to the funeral

Were the new shoes from the old lady appropriate for a funeral?
a) Strongly certain the new shoes from the old lady were NOT appropriate for a funeral
b) *Strongly certain the new shoes from the old lady were appropriate for a funeral