Reading Comprehension, Bridging Inferences, and their Relation to Working Memory Processes in Children in Grades Three through Eight

by

Amber Michele Johnston

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctorate of Philosophy
in
Psychology

Guelph, Ontario, Canada

ABSTRACT

READING COMPREHENSION, BRIDGING INFERENCES, AND THEIR RELATION TO WORKING MEMORY PROCESSES IN CHILDREN IN GRADES THREE THROUGH EIGHT

Amber Michele Johnston
University of Guelph, 2012

Advisors:
Professor M. A. Barnes
Professor R. W. Barron

Reading comprehension (RC) is a complex, dynamic process by which readers interact with text to construct meaning. It relies on word-level decoding and vocabulary skills, text-level skills such as inference, and general-purpose cognitive processes such as working memory (WM). Bridging inferences, which connect different parts of text to maintain semantic coherence, are necessary for comprehension. WM is thought to provide the mental workspace that allows readers to comprehend text, including making bridging inferences. This dissertation includes two studies that investigate related but unique questions regarding inference-making, WM, and RC in school aged children. The first study demonstrated that bridging inference making increased with age and was affected by text distance such that inferences across larger chunks of text were more difficult to make than those between adjacent sentences. Bridging inferences were also affected by knowledge domain such that affective inferences based on character goals, desires, or motivations were made correctly less often than were concrete inferences based on concrete, causal information. Semantic updating, an aspect of WM that involves efficiently revising the contents of WM, predicted variance in the far-concrete inferencing condition. Semantic reactivation, an aspect of WM that involves bringing previously processed information back into WM, predicted variance in the far inferencing conditions regardless of knowledge domain. The second study investigated the word-level and text-level skills and general purpose cognitive processes that support performance on three different standardized RC measures. Semantic updating and semantic reactivation predicted variance on the RC tool considered to measure mental model *building* (WJIII-PC, Woodcock-Johnson-III passage comprehension subtest). Semantic reactivation also predicted variance on the RC tool considered to measure the ability to build and then *reflect upon* the mental model (WIAT-II-RC, Wechsler Individual Test of Achievement reading comprehension subtest). A measure of verbal WM predicted performance on one RC measure in the presence of word-level and text-level skills but only when the hypothesized components of WM (updating and reactivation) were not included in the model. Additionally, bridging inference making uniquely predicted performance on all three RC measures. The findings suggest readers coordinate different cognitive processes depending on the comprehension task.

Acknowledgements

I would like to thank Dr. Harvey Marmurek for chairing my oral defense and for his helpful and enthusiastic comments regarding my dissertation. Thank you to Dr. Mary Ann Evans for agreeing to sit on my defense committee, for her thoughtful questions, and for her thorough evaluation of my dissertation. And thank you to Dr. John Kirby for agreeing to be the external committee member of my defense committee, and for his expertise and thought provoking questions.

I would like to thank my supervisor, Dr. Marcia Barnes, for her time and dedication to this project. Marcia, with your support I was able to fulfill a dream and find the courage and confidence to follow my own path. Your enthusiasm and dedication to research and teaching are incredible and I hope my work will always be driven by the kind of passion you have. Thank you for joining me on this part of my journey – I hope our paths cross again.

Thank you to my co-supervisor, Dr. Rod Barron, for his guidance and expertise throughout my entire graduate career. Rod, you taught the very first and very last graduate courses I completed at the University of Guelph; you have always encouraged me to stretch farther than I thought possible in my learning. I deeply appreciate your devotion to students and sincerely admire your excitement about research.

Although the following women did not directly support this research project, they are among the accomplished and kindhearted women who have mentored me throughout my academic and clinical career. I feel privileged to have learned from each of you. Thank you Dr. Margo Watt, Dr. Brenda Kenyon, Ms. Eileen Gross, and Dr. Julie Perkins!

I would also like to thank my classmates who brought laughter when it was needed the most, encouragement when I was sure I would give up, and who allowed me to borrow their hope when I could not find my own. Dr. Cailey Hartwick, your friendship and wisdom are immeasurable; Dr. Nora

Klemencic, your devotion and compassion are an inspiration; Dr. Michelle Wesley, your optimism and courage are contagious. I will be forever grateful for you, my fellow Blacksheep – I would be honored to be "unherdable cats" with you girls any day! To the Barnes lab girls, Landa (Fox) Mark, Jori Jones, Meredith Pike, and Brownen Davis, thanks for never actually kicking me out for bringing stinky eggs into our lab each morning and for joining me on so many drives back and forth to Hamilton.

To my Mom and Dad, Joan and Ray Johnston, thank you for always believing in me, especially when I did not believe in myself. By watching both of you I learned that with passion and hard work I could make my dreams come true. You celebrated each and every success with me, and were always there to offer words of wisdom and a safe place to rest my head when life got tough. I will be forever grateful that I have you both in my life.

To Mike Bridle, you have never known me without a thesis attached to my heel. I will never be able to put into words how much your love and support have lifted me throughout the past eight years – I could not be more excited to take the next step in my life with you by my side. Get ready, 'cause we have big plans ahead of us, Babe!

~ It is not because things are difficult that we do not dare,
rather it is because we do not dare that they are difficult.

Table of Contents

List of Tables

Chapter 1: General Introduction

Reading Comprehension is a dynamic process through which readers interact with the text to construct meaning. Word-level skills, such as word decoding and vocabulary knowledge, text-level skills, such as bridging inferences, and general cognitive processes, such as working memory, support readers in the comprehension process. This dissertation includes two related but distinct studies that investigate questions about the process of reading comprehension in school aged children. Study 1 is in Chapter 2; Study 2 is in Chapter 3. Some of the measures and information about participants are common to both studies. For ease of presentation, details on overlapping measures and participants are provided only in Study 1.

Overview of Study 1

Bridging Inferences: Text Distance and Domain Knowledge

The first cross-sectional study, *Bridging Inferences and their Relations to Working Memory Processes in Children in Grades Three through Eight*, investigates the development of bridging inferences across grades, the effects of text distance and domain knowledge, and which working memory components (semantic updating and semantic reactivation) support bridging inferences. It is well established that the ability to make bridging inferences increases with age (Barnes, Dennis, & Haefele-Kalvaitis, 1996; Casteel & Simpson, 1991; Johnston, Barnes, & Desrochers, 2008; Pike, Barnes, & Barron, 2010). Making bridging inferences between information that is presented close together in the text is easier for typically developing children compared to making bridging inferences across larger chunks of text (Johnston, et al., 2008). Studies of children with age appropriate decoding skills but poor comprehension have found that although all children are less skilled when making inferences across large textual distances, increased distances affect poor comprehenders to a greater extent (Barnes, Faulkner, Wilkinson, & Dennis, 2004; Cain, Oakhill, & Elbro, 2003; Cain, Oakhill, & Lemmon, 2004). The current study is the first to examine the effects of knowledge domain on making bridging inferences in general as well as across two text distances. The experimental bridging inference task included items that reflected concrete, causal information (concrete inferences) and those that reflected character goals, desires, or motivations (affective inferences). Previous research with university students suggests that making inferences with concrete, causal information is less process intensive than making inferences based on character goals/plans (Shears, Miller, Ball, Hawkins, Griggs, & Varner, 2007).

Bridging Inferences, Semantic Updating, and Semantic Reactivation

Working memory has been linked to inferencing in many studies (Cain et al., 2004; Johnston et al., 2008; Pike et al., 2010). This is the first study to look at the cognitive resources needed to make bridging inferences as a function of text distance (near versus far) and domain knowledge (concrete versus affective). Specifically, this study investigated the relations between specific working memory processes thought to be most important for comprehension (semantic updating and semantic reactivation) and bridging inferences. Semantic updating, which involves revising the contents of working memory as new and relevant information is presented (Carretti, Cornoldi, De Beni, & Romano, 2005), is hypothesized to be important for integrating pieces of information together to make inferences and construct meaning. Semantic reactivation, the ability to bring previously processed information back into working memory from long-term memory (O'Brien, Cook, & Gueraud, 2010), is hypothesized to be important for making inferences particularly in the far bridging inference condition where the to-be-integrated information is separated by chunks of text (several sentences). Accordingly, the aims of this study included investigations of the following: grade-related changes in bridging inference making; the effects of manipulating text distance and knowledge domain on bridging inference making; and the relation between bridging inferences and specific working memory processes (semantic updating and semantic reactivation) in school aged children.

Overview of Study 2

What do Reading Comprehension Tests Measure?

The second study, *The Relation of Working Memory Processes and Bridging Inferences to Reading Comprehension in Children in Grades Three through Eight*, investigated the relations of working memory processes (semantic updating and semantic reactivation) to performance on three commonly used, standardized reading comprehension measures - Passage Comprehension subtest from the Woodcock-Johnson III Test of Achievement, WJIII-PC (Woodcock & Johnson, 2001), Paragraph Reading subtest from the Test of Reading Comprehension, third edition (TORC3-PR, Brown, Hammill, & Wiederholt, 1995), and Reading Comprehension subtest from the Woodcock-Johnson Individual Achievement Test, second edition (WIAT-II-RC, Wechsler, 2002). The skills and abilities needed to demonstrate adequate comprehension skill may vary depending on the measure used to assess a complex and multi-faceted skill such as reading comprehension (e.g., Bowyer-Crane & Snowling, 2005; Cutting & Scarborough, 2006; Francis, Fletcher, Catts, & Tomblin, 2005; Francis, Snow, August, Carlson, & Iglesias, 2006; Johnston, 2006; Keenan, Betjemann, & Olson, 2008; Nation & Snowling, 1997; van den Broek, Rapp, & Keough, 2005). Keenan et al. (2008) suggested that shorter passages are more reliant on word-level skills such as decoding because there is less information presented in the text to support gap-filling, such as using context to infer the meaning of a word when one is unable to decode it. A text-level skill, such as bridging inference making, has been related to a variety of reading comprehension measures across several studies (Johnston et al., 2008; Pike et al., 2010; Schmalhafer, McDaniel, & Keefe, 2002).

Reading Comprehension, Working Memory, and Working Memory Components

Working memory predicts reading comprehension (Cain et al., 2004) and researchers have suggested that the relation between performance on reading comprehension and working memory tasks is a reflection of the relation of specific components of working memory with comprehension (Christopher, Miyake, Keenan, Pennington, DeFries, Wadsworth, Willcutt & Olson, 2012). Christopher et al. investigated the unique relation of working memory and working memory components that consisted of suppression/inhibition, processing speed, and word naming speed with reading comprehension and word reading outcome measures. Their findings demonstrated that both working memory and processing speed predicted unique variance in word reading and reading comprehension, highlighting the importance of being able to hold and manipulate information in memory and quickly process visual information for both word reading and reading comprehension.

The working memory components studied in study one, semantic updating and semantic reactivation, have been linked to reading comprehension (Gernsbacher, 1990; Palladino, Cornoldi, DeBeni, & Pazzaglia, 2001). However, these two hypothesized comprehension-related aspects of working memory have not been studied together in children's reading comprehension, nor has their relation to different reading comprehension measures been assessed. Three standardized reading comprehension measures were chosen because the passages included in each measure varied in length and because each of the three measures used different response formats (cloze, multiple choice questions, and open ended questions). Both factors have been argued to underlie the performance differences among reading comprehension measures (Keenan et al., 2008; Nation & Snowling, 1997). Additionally, the WJIII-PC and TORC3-PR were related to different working memory and attentional processes in a previous study (Johnston, 2006). The aim of the second study was to investigate the relations of word-level and text-level skills investigated in study one (bridging inferences) and those working memory processes hypothesized to be implicated in reading comprehension (semantic updating and semantic reactivation) to performance on three different standardized reading comprehension measures (e.g., WJIII-PC; TORC-3; WIAT-II-RC) in children ranging from grades three through eight

References

Barnes, M. A., Dennis, M., & Haefele-Kalvaitis, J. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inferencing in children from six to fifteen years of age. *Journal of Experimental Child Psychology, 61*(3), 216-241.

Barnes, M. A., Faulkner, H. Wilkinson, M., & Dennis, M. (2004). Meaning construction and integration in children with hydrocephalus. *Brain and Language, 89*, 47 - 56.

Castell, M. A., & Simpson, G. B. (1991). Textual coherence and the development of inferential generation skills. *Journal of Research in Reading, 14*(2), 116-129.

Brown, V., Hammill, D. D., & Wiederholt, J. L. (1995).*The Test of Reading Comprehension-Third Edition*. Austin, TX: PRO-ED.

Bowyer-Crane, C., & Snowling, M. J. (2005). Assessing children's inference generation: What do tests of reading comprehension measure? *British Journal of Educational Psychology, 75*, 189 – 201.

Cain, K., Oakhill, J. V., &Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language, 30*, 681 - 694.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96*, 671–681.

Carretti, B., Cornoldi, C., De Beni, R., & Romano, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology, 91*, 45 – 66.

Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J., Willcutt, E., & Olson, R. K. (2012, February 20). Predicting Word Reading and Comprehension With Executive Function and Speed Measures Across Development: A Latent Variable Analysis. *Journal of Experimental Psychology: General*, *141*(3), 470 – 488.

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10*, 277 - 299.

Francis, D. J., Fletcher, J. M., Catts, H. W., & Tomblin. (2005). In *Children's reading comprehension and assessment*. Eds. Paris, Scott G., Stahl, Steven A. (369 – 394. Mahwah, NJ, US: Lawrence Erlbaum.

Francis, D. J., Snow, C. E., August, D., Carlson, C. D., Miller, J., & Iglesias, A. (2006). Measures of reading comprehension: A latent variable analysis of the diagnostic assessment of reading comprehension. *Scientific Studies of Reading, 10*(3), 301 – 322.

Gernsbacher, M. A. (1990).*Language comprehension as structure building.* Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Johnston, A. M. (2006).*The role of attentional control in reading comprehension*. Unpublished master's thesis, University of Guelph, Guelph, Ontario, Canada.

Johnston, A. M., Barnes, M. A., &Desrochers, A. (2008, June). Bridging Inferences in School Aged Children. Poster presented at the 39th Annual Convention of the Canadian Psychology Association (CPA), Halifax, NS.

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*(3), 281 - 300.

Nation, K., & Snowling, M. (1997). Assessing reading difficulties: the validity and utility of current measures of reading skill. *British Journal of Educational Psychology, 67*, 359–370.

Palladino, P., Cornoldi, C., DeBeni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory & Cognition, 29*(2), 344-354.

Pike, M. M., Barnes, M. A., & Barron, R. W. (2010). The role of illustrations in children's inferential comprehension. *Journal of Experimental child Psychology, 105*, 342 – 255.

Schmalhofer, F., McDaniel, M. A., & Keefe, D. (2002). A unified model for predictive and bridging inferences. *Discourse Processes*, 33(2), 105-132.

Shears, C., Miller, V., Ball, M., Hawkins, A., Griggs, J., & Varner, A. (2007). Cognitive demand differences in causal inferences: Characters' plans are more difficult to comprehend than physical causation. *Discourse Processes, 43*(3), 255 – 278.

van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes, 39*(2 & 3), 299 - 316.

Wechsler, D. (2002). *Manual for the Wechsler Individual Achievement Scale, 2nd Edition.* Toronto: The Psychological Corporation.

Woodcock, R., Mather, N., & McGrew (2001).*Woodcock-Johnson III – Test of Achievement.* Chicago: Riverside.

Chapter 2: Bridging Inferences and their Relation to Working Memory Processes in Children in Grades Three through Eight

Bridging inferences are implicated in reading comprehension because they serve to integrate explicitly stated ideas within text that may be necessary for deriving meaning (Cain & Oakhill, 1999; Kintsch & Kintsch, 2005; Long, Oppy, & Seely, 1997). Bridging inferences can be made within sentences or across larger chunks of text such as sentences or paragraphs. If those connections that are important for specifying the meaning of the text are not made, comprehension will suffer (Barnes, Faulkner, Wilkinson, & Dennis, 2004; Barnes, Huber, Johnston, & Dennis, 2007; Long et al., 1997). Skilled adult readers often make inferences in order to maintain coherence (van den Broek, 1994) and older children tend to make inferences more often and more accurately than younger children (Barnes, Dennis, & Haefele-Kalvaitis, 1996; Casteel & Simpson, 1991; Johnston, Barnes, & Desrochers, 2008; Pike, Barnes, & Barron, 2010).

Concrete and Affective Bridging Inference

van den Broek, Kendeou, Kremer, Lynch, Butler, & Lorch (2005) proposed that children progress through a trajectory of inferential development in which they first make inferences that are more concrete, such as those based on physical cause and effect relations (hereafter referred to as concrete inferences). With increasing age, they become more consistent in making bridging inferences with information involving emotional states, desires, goals, and motivations of the characters described in the text. Readers monitor and track abstract information, such as the goals of protagonists (Linderholm, Gernsbacher, van den Broek, Neninde, Robertson, & Sundermier, 2004), to help construct meaning while reading. Readers can use information about character goals and the situations that characters are involved in to make inferences about characters' emotional states (Molinari, Burin, Saux, Barreyro, Irrazabal, Bechis, Duarte, & Ramenzoni, 2009). However, even adult readers are faster (Campion & Rossi, 2001; Singer, Halldorson, Lear, & Andrusiak, 1992) and more accurate (Shears & Chiarello, 2004; Shears, Miller, Ball, Hawkins, Griggs, & Varner, 2007) when making inferences based on knowledge about physical causation (e.g., inferring that pouring water on a fire caused the fire to go out after reading “Zander poured the bucket of water on the small campfire. The fire went out”) compared to knowledge about character-based goals and plans (e.g., inferring that Joan was likely shopping for a gift to bring to the party after reading “Joan left early for the birthday party. She spent an hour shopping at the mall”). For the purposes of this paper, these latter types of inferences will be referred to as affective inferences.

Text distance

Text distance affects the making of inferences (van den Broek et al., 2005) because inferences are made more often and consistently when the information needing integration is presented in close proximity rather than farther apart in the text. Studies that manipulate the conditions under which bridging inferences are made have found that children are less accurate (Cain, Oakhill, &Elbro, 2003, Johnston et al., 2007) and slower (Barnes et al., 2004) when integrating information across larger textual distances than across shorter distances. Children who are skilled decoders but who struggle to understand what they read (poor comprehenders) are more negatively affected by text distance compared to skilled comprehenders (Barnes et al., 2004; Cain et al., 2003). In some studies, skilled and less skilled comprehenders are comparable in their inference performance when the information to-be-integrated is close together in the text (Cain et al., 2003). In a cross-sectional study of bridging inferences (Johnston, Barnes, & Davis, 2007), children of all ages appeared to be equally disadvantaged by larger textual distances.

Inferential Comprehension and Verbal Working Memory

There is considerable evidence that working memory is important for making bridging inferences (Cain et al. 2004; Johnston et al., 2008; Pike et al., 2010). Cain et al. (2004) found that children who performed more poorly on a task that required them to infer new word meanings from context across shorter and longer text distances also had lower scores on a verbal working memory task. Cain and colleagues proposed that when the distance between information needing to be integrated was greater, the processing demands on working memory increased. Pike et al. (2010) also found a link between bridging inferences and working memory. In their study of bridging inferences and illustrations, verbal working memory accounted for unique variance when information had to be integrated across several sentences in a text even after controlling for word reading skill and vocabulary knowledge.

Very little research has been conducted that looks at the cognitive resources needed to make inferences using different types of knowledge (e.g., cause and effect or concrete versus character-based or affective). Using a dual task procedure, Shears et al. (2007) found that the number of inferences university students were able to make while performing a dual memory task (i.e., working memory) differed depending on the information used to make the inference. Inferences based on character goals/plans were more process intensive than those based on cause and effect relations.

To the author's knowledge, no study has explored whether working memory differentially predicts bridging inference-making as a function of text distance (near versus far) and type of information used (concrete versus affective). Nor has any study sought to investigate what aspects of working memory, such as updating or reactivation processes, might be most related to the making of bridging inferences in children.

Working Memory Processes: Semantic Updating and Semantic Reactivation

Working memory is a mental workspace that has a limited capacity (Unsworth & Engle, 2007). It includes components or sub-processes that must be used strategically in order to efficiently perform complex tasks, such as comprehending text (Baddeley, 1986; Unsworth & Engle, 2007), which includes making inferences. Particular working memory components have been hypothesized to be associated with comprehension in general, such as suppression/inhibition, processing speed, and naming speed (Christopher, Miyake, Keenan, Pennington, DeFries, Wadsworth, Willcutt & Olson, 2012) and inference making in particular, such as updating (Barnes et al., 2007) and reactivation and suppression (Barnes et al., 2004; Gernsbacher, 1997; Gernsbacher & Shlesinger, 1997; Pimperton & Nation, 2010). The process of updating, which involves efficiently revising the components of working memory as new and relevant information becomes available (Carretti, Cornoldi, De Beni, & Romano, 2005), may be important for integrating new pieces of information from the text with previously read text to form coherent mental models. During the comprehension process, information that is no longer relevant is thought to be suppressed or inhibited to prevent it from being integrated into the mental model and interfering with comprehension (Barnes et al., 2004; 2007; Gernsbacher, 1997; Gernsbacher & Faust, 1991; Gernsbacher & Shlesinger, 1997; Pimperton & Nation, 2010). As the reader progresses through the text, information that was previously processed but no longer held in working memory must be reactivated or brought back into working memory from long term memory (O'Brien, Cook, & Gueraud, 2010; O'Brien, Rizzella, Albrecht, & Halleran, 1998; Rapp & van den Broek, 2005) to help readers construct meaning (including making bridging inferences). Reactivation should be essential for making bridging inferences particularly when the information to-be-integrated is separated by large chunks of text. The current study seeks to clarify the relations between semantic updating and semantic reactivation and the making of bridging inferences under different text distance conditions and in different knowledge domains.

Aims and hypotheses.

The present study tested the influence of knowledge domain and text distance on bridging inference-making in school-aged children. The revised Bridging Inferences Task (Bridge-IT) used by Johnston et al. (2007) includes items that tap cause and effect relations (concrete inferences) and those related to character goals, desires, and motivations (affective inferences). Readers were required to make concrete or affective inferences across two different textual distances (near versus far). This study tested whether textual distance (near versus far) and knowledge domain (concrete versus affective) affect the making of bridging inferences across grades three through eight, and also whether different working memory processes (updating and reactivation) are related to bridging inference performance in each condition.

The first aim was to explore the grade related effects of bridging inferences and the type of errors that characterize bridging inferences in typically developing school-aged children in grades three through eight. Guided by previous findings (Barnes et al., 1996; Johnston et al., 2007; Pike et al., 2010), it was hypothesized that younger children would be less accurate when making bridging inferences compared to older children.

The second aim was to investigate the effects of manipulating text distance used to make bridging inferences. Based on previous findings (Barnes et al., 2004; Cain et al., 2003; Johnston et al., 2007) it was hypothesized that children would be less accurate when making bridging inferences over larger versus shorter text distances (Bridge-IT-far versus Bridge-IT-near conditions). The types of errors that children made on the Bridge-IT also allowed for an indirect investigation of the role of the suppression processes in inference making. The Bridge-IT required children to read a five-sentence paragraph containing two related but competing mental models and then select one of three sentences to complete the passage (see Tables 1.1a and 1.1b for examples). Choosing the best sentence to complete the story suggested that children made the correct bridging inference. Of the other two (incorrect) sentences, one reflected the competing mental model while the other was related to both models but was highly unlikely given the context of the passage. It was hypothesized that when children choose the incorrect sentence that reflects the competing mental model, it is because they are unable to suppress or ignore the irrelevant, competing information or because they are unable to retrieve the necessary information presented in the text, or both. In addition, it was thought that the suppression of irrelevant information (Cain et al., 2004) and/or the reactivation or retrieval of prior relevant information from the text might be most difficult in the far condition and, therefore, there would be more errors in the far condition that reflect lack of suppression compared to errors due to choosing an improbable, but contextually-related response. A more detailed explanation about why suppression and/or reactivation may be important for performance on the Bridge-IT is illustrated with an example in the Methods section.

The third aim was to investigate the effects of manipulating domain knowledge on bridging inferences. Based on studies that suggest making inferences with causal knowledge develops earlier than inferences using affective knowledge and also that the former type of inference is made more easily than the latter even in adults (Shears & Chiarello, 2004), it was also hypothesized that children would find it more difficult to make affective inferences compared to concrete, causal inferences.

The fourth aim was to explore the relation between bridging inferences and specific working memory processes, namely semantic updating and semantic reactivation. It was hypothesized that updating would be important for bridging inferences in all conditions because maintaining coherent mental models requires updating the contents of working memory as new information is presented. Reactivation was predicted to be most important in the far inferencing condition because children would need to reactivate or retrieve the first mental model (target sentence) in order to make the correct inference. Based on Shears & Chiarello's (2004) findings, one research question was whether the pattern of relations between working memory components and affective inferences might differ from the relations between working memory components and concrete inferences.

Method

Participants

One hundred and thirty-eight English-speaking children (70 boys and 68 girls) ranging from grades three through eight (8- to 14-years-old), were recruited from a local school board in south-central Ontario to participate in the present study. Children were excluded from participating in the study if they were enrolled in a special education class or had been identified with a learning disability; had a gestational age less than 36 weeks; had major neurological or behavioural disorders (e.g., head injuries with hospitalization, autism, disruptive behaviour disorder), or were being schooled in English as a second language. Additionally, children were required to have grade-appropriate word reading skills as assessed by the *Sight Word Efficiency subtest from the Test of Word Reading Efficiency* (TOWRE; Torgensen, Wagner, & Rashotte, 1999). Participants were excluded if their score was one or more standard deviations below the population mean for their age. Two students with below average scores on this measure of word decoding were excluded from the study based on initial screening; one child was excluded because English was his second language; six students were excluded because they were identified by parents has having a learning disability (e.g., reading disability; mathematics disability); and one grade seven student was removed due to a missing score (word decoding). The final sample consisted of 128 children in total that included 19 grade three students (*M* age = 8.21), 25 grade four students (*M* age = 9.16), 23 grade five students (*M* age = 10.39), 21 grade six students (*M* age = 11.38), 21 grade seven students (*M* age = 12.38), and 19 grade eight students (*M* age = 13.21).

Measures

Sight Word Efficiency subtest from the Test of Word Reading Efficiency (TOWRE; Torgensen et al., 1999). The TOWRE was used to obtain a measure of word reading efficiency, which involved accuracy and fluency. This commonly used test required each child to read as many words (104 words in total) as they could in the allotted time (45 seconds). The words became progressively more difficult. Although standardized scores were used in the screening process, raw scores were used in analyses because experimental tasks measuring working memory components (e.g., semantic updating, semantic reactivation) were also used in the analyses. American test norms were used to derive standardized scores. The test-retest reliability is $r = .97$ for children aged six to nine years and $r = .84$ for children aged 10 to 18 years.

The Picture Vocabulary subtest from the Woodcock-Johnson III (WJIII-voc; Woodcock et al., 2001). The WJIII-voc was used to obtain a measure of vocabulary knowledge. Each child was presented with a page including six pictures and was asked to name each picture using one word. Raw scores were used in analyses for the same reasons mentioned above. American test norms were used to derive standardized scores. This test includes norms ranging from 2 years to over 90 years of age and the test-retest reliability is $r = .82$ for children between the ages of 2 to 18 years when tested one to two years later.

Bridging Inferences Task (Bridge-IT). The Bridge-IT, an experimental task, measures the ability to integrate and make inferences using information presented within text. The text was manipulated in two ways: distance and type of information presented. Information needing to be integrated was either in close in proximity in the text (i.e., contiguous sentences) or farther apart in the text (i.e., separated by four sentences). The type of inferences made in each distance condition (i.e., near or far) required making inferences with cause and effect concrete information (e.g., it is unlikely that someone with a broken leg will compete in a jumping rope competition; see Table 1.1a for example) or affective information (e.g., it is likely that someone who is frightened of sharks and other fish that swim in the ocean would be cautious about swimming in the ocean; see Table 1.1b for example). Each child completed five near-concrete inference items, five far-concrete items, five near-affective inference items, and five far-affective inference items.

Each child read 20 short stories (containing five sentences each) and was instructed to choose "the next best sentence" from three possible answers "to complete the story." The position of answers (e.g., a, b, or c) were counter balanced so that equal numbers of near and far inferences were made from either the first or last sentence to answers a, b, or c. Each child was given a 20-page booklet. The cover page included detailed directions, which were read to each child. Upon turning the first page, the child was presented with the first five-sentence story. The child was instructed to turn the page when s/he finished reading each story. On the backside of each page were three answer choices. The child was asked to choose the best sentence to come next in the story and was not allowed to turn back to refer to the story s/he had just read. The child progressed through the booklet at his or her own pace. In the near condition, the target sentence was always the last sentence in the story and included the target concrete or affective information. In the far condition, the target sentence was always the first sentence in the story and included the target concrete or affective information. Choosing the best of the three possible answers demonstrated that the child made the correct inference.

Table 1.1a

Example of causal-near and -far conditions from the Bridging Inferences Task (Bridge-IT)

Causal-near condition: BRIDGE-IT-near:

Jenny is at school and it is time for gym class.

Jenny enjoys playing lots of games in gym.

Today in gym, her class is having the big skipping rope competition.

Jenny loves skipping and last year she won the skipping competition.

Jenny has a broken leg from falling down the stairs last week.

a) *This time Jenny helped judge the skipping competition.*

b) This time Jenny won the skipping competition again.

c) This time Jenny left class and ran away from school.

Causal-far condition: BRIDGE-IT-far:

Jenny has a broken leg from falling down the stairs last week.

Jenny is at school and it is time for gym class.

Jenny enjoys playing lots of games in gym.

Today in gym, her class is having the big skipping rope competition.

Jenny loves skipping and last year she won the skipping competition.

a) *This time Jenny helped judge the skipping competition.*

b) This time Jenny won the skipping competition again.

c) This time Jenny left class and ran away from school.

Note. Target sentences are in bolded font; correct answers are in italicized font; choice corresponding to competing model is in underlined.

Table 1.1b

Example of affective-near and –far conditions from the Bridging Inferences Task (Bridge-IT)

Affective-near condition:

Leanne loves to swim and is becoming a lifeguard at the pool this summer.

Leanne went with her friends to the beach for a day of sun tanning and swimming.

It is a very hot day today, and the water looks cool and refreshing.

All of Leanne's friends ran into the water as soon as they put their towels down.

Leanne is scared of the sharks and other big fish that swim in the ocean.

a) Leanne picks a bouquet of flowers.

b) Leanne runs into the water.

c) *Leanne hesitantly wades in up to her knees.*

Affective-far condition:

Leanne is scared of the sharks and other big fish that swim in the ocean.

Leanne loves to swim and is becoming a lifeguard at the pool this summer.

Leanne went with her friends to the beach for a day of sun tanning and swimming.

It is a very hot day today, and the water looks cool and refreshing.

All of Leanne's friends ran into the water as soon as they put their towels down.

a) Leanne picks a bouquet of flowers.

b) Leanne runs into the water.

c) *Leanne hesitantly wades in up to her knees.*

Note. Target sentences are in bolded font; correct answers are in italicized font, choice corresponding to competing model is underlined.

The rationale behind this measure is based on reading comprehension models (Gernsbacher, 1990). Each story consists of two opposing mental models within the first five sentences (refer to Table 1.1a for example). It is presumed that the reader must suppress or ignore one model (the child loves skipping and is really good at it) and/or retrieve previously read information (she has injured her leg) in order to integrate the target model (that she will have trouble performing physical activities) with one of the three choices, resulting in the correct inference (she helped judge the skipping competition instead of participating). In the near condition, the second model presented was the target model and was to be integrated with one of the three choices. In the far condition, the target model was presented first, leaving the need for the second model to be suppressed or ignored and/or the first model to be reactivated in order to make the correct inference.

One point was awarded for each correct answer. Each condition (Bridge-IT-near-concrete, Bridge-IT-near-affective, Bridge-IT-far-concrete, and Bridge-IT-far-affective) was scored out of a possible 5 correct.

The types of errors that children made were also coded and recorded. If the child chose an incorrect answer that corresponded to the competing model (the to-be-ignored or suppressed model) then a processing error (PE) was recorded (she won the skipping competition). If the child chose an incorrect sentence that was improbable, an improbable response error (IRE) was recorded (she left class and ran away from school).

The Bridge-IT included two versions. Each version contains identical items with the only difference being textual distance. For example, if in version A the skipping item requires that a near inference be made, the skipping item on version B required that a far inference be made. Children within each grade were randomized to receive either version A or B. The parallel measures reliability coefficient for the Bridge-IT-global was 0.66 for version A and 0.71 for version B (Field, 2005).

The Semantic Reactivation Task. This task was designed to measure the ability to reactivate or retrieve previously processed information when cued to do so. Each child was asked to read lists of words. The words were from two semantic categories (e.g., occupation and fruit; see Table 1.2 for example). Immediately after reading the word list, the page was turned by the examiner to reveal an action picture (e.g., children splashing in a puddle). The child was asked, "What word immediately comes to mind to describe this picture?" After providing one word that described the picture, the page was turned and a blank page was revealed. The examiner immediately cued the child to recall all the words from one semantic category (e.g., fruit) by saying the last word belonging to that semantic category (e.g., banana). For example, if the target category was fruit and the non-target category was occupation for the following list: teacher, policeman, plum, doctor, apple, banana, after the child named the picture, the examiner would say, "banana, what word(s) from the list you just read go with banana?" The correct response would be: plum and apple, in any order.

The presentation and naming of the picture served to prevent rehearsal of the word list and switch the child's attention from the word list just read so that the words would be pushed out of the focus of attention, providing the condition under which the words must be brought back into the focus of attention or reactivated when cued to do so. Cronbach's alpha reliability coefficient was .75 for the reactivation task.

Table 1.2

Example of Semantic Reactivation Task

List read by child	Describe the picture (child responds)	Cued recall (stated by examiner)	Correct recall (by child)
teacher policeman **plum** doctor **apple** *banana*		*banana*	**apple** **plum**

Note. Target words are in bolded font; cue word in italicized font.

The task demands varied across two dimensions: memory load and proactive interference. The child had to recall 1, 2, 3 or 4 words from a list of 4, 6, 8, or 10 words in total, in three different proactive interference conditions. The child was asked to recall words from a semantic category that: 1) was entirely new (no interference condition), 2) had been a non-target or distractor category in a previous item (low proactive interference), or 3) had been a target category in a previous item (high proactive interference). Each item contained new words for the semantic category, regardless of whether the semantic category had been encountered in a previous item or not. Each child completed six items from each memory condition. Two items contained no interference, two items contained low proactive interference, and two items contained high proactive interference. One point was awarded for each correctly recalled word. The total number of words that could be recalled across the entire task was 60 words (see Appendix A for grade-related findings). Because the amount of proactive interference (or lack of) encountered while reading may vary depending on text distance and domain knowledge, the total raw score was used in all analyses in an attempt to capture the reactivation processes that may be engaged while reading passages from different conditions across the bridging inference task.

The Semantic Updating Task. This task was based on the Palladino, Cornoldi, De Beni, and Pazzaglia's (2001) updating task. It was designed to measure the ability to update information according to semantic content (i.e., ordering according to size), an important working memory process (Garson, Bryson, & Smith, 2008; Palladino et al., 2001). Each child was assessed individually. The examiner read a 10-word list at the rate of one word per second that included animal words and words that are not animals (abstract distracter words, such as hope and strength). There were two memory loads: the child was asked to recall either the three smallest or the four smallest animal words from each list. The number of updates the child needed to make in order to correctly recall the three (or four) smallest animal words varied (either five or six updates). To illustrate, the child must hold the first three animal words heard from the list, then upon hearing an animal that is smaller than the smallest animal held in memory, the child must drop the largest animal held in memory and update the list to include the new, smaller animal. Each time a child added an animal or revised the information in working memory, an update was assumed (see Table 1.3 for an example). The raw score was calculated by adding the number of target words recalled in the correct order, from smallest to largest, for each item. The total number of words that could have been recalled was 20 across 6 trials (see Appendix A for grade-related findings). Cronbach's alpha reliability coefficient was .59 for the updating task. Because the number of updates made while forming mental models while reading may vary depending on the demands of the text being read, such as text distance and domain knowledge, the total raw score was used in all analyses in an attempt to capture the updating processes that may be required across different conditions on the Bridge-IT.

Table 1.3

Example of a 3 Item Trial with 5 Updates from the Semantic Updating Task

Word list	Presumed cognitive process
bear	add to working memory (first update to information held in working memory)
skunk	add to working memory (and update order)
energy	Distracter (ignore or suppress)
leopard	add to working memory (and update order)
strength	distracter (ignore or suppress)
worm	add to working memory (update order, remove or suppress bear)
ox	Remove from memory or suppress
faith	distracter (ignore or suppress)
kitten	add to working memory (update order, remove or suppress leopard)
seal	Remove from memory or suppress

Note. Target words are in bolded font. Correct recall: worm, kitten, skunk.

Procedure

Children who passed the word reading screening measure were tested individually on measures of vocabulary knowledge, verbal working memory, semantic reactivation and updating, and tested in small groups on the measure of bridging inferences, in 45-60 minute blocks over the course of several sessions.

Results

Participant characteristics and scores on standardized measures of vocabulary knowledge and word decoding skills are in Table 1.4. To test if there were differences in word reading skill or vocabulary knowledge between grades, one-way ANOVAs were conducted on each of the standardized measures using standardized scores. Main effects of grade were found for word reading (F (5, 122) = 2.98; p< 0.05) and vocabulary knowledge (F (5, 122) = 3.20; p<0.01). Tukey post-hoc tests (p< .05) revealed that the grade three word reading standard scores were significantly higher than the standard scores of grades five scores. In terms of vocabulary standard scores, the grade three scores were significantly higher than the scores of grades five and seven.

Table 1.4

Means (standard deviations) for standardized measures of vocabulary and word reading

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
n	19	25	23	21	21	19
TOWRE-raw	65.74	68.36	70.52	78.33	81.71	85.68
	(5.55)	(8.44)	(9.04)	(8.95)	(8.73)	(7.76)
TOWRE-ss	112.79	108.56	101.43	105.86	105.05	105.95
	(7.43)	(10.10)	(10.32)	(11.13)	(10.87)	(9.76)
WJIII-voc-raw	24.68	25.32	26.13	28.43	28.86	31.95
	(1.86)	(2.01)	(2.75)	(2.82)	(2.52)	(2.84)
WJIII-voc-ss	110.37	105.76	103.09	104.81	102.86	109.53
	(7.85)	(7.19)	(7.85)	(9.07)	(7.42)	(8.84)

Note. TOWRE-raw = raw scores from the Sight Word Efficiency subtest of the Test of Word Reading Efficiency (TOWRE);

TOWRE-ss = standard scores from the Sight Word Efficiency subtest of the TOWRE; WJIII-voc-raw = raw scores from the Picture Vocabulary subtest of the Woodcock-Johnson test of Achievement (WJIII); WJIII-voc-ss = standard scores from the Picture Vocabulary subtest of the WJIII.

Bridging Inferences

To test the effects of grade, distance, and knowledge domain on inferencing accuracy, a six group (grades three through eight) by two-integration distance (near versus far) by knowledge domain (concrete versus affective), mixed analysis of variance (ANOVA) was conducted. The analysis revealed a main effect of grade (F (5, 122) = 2.36; $p < .05$; η^2_p: .09). Dunnett t-tests (using grade 8 as control mean) revealed that the youngest children (grades three and four) were less accurate than the oldest children (grade eight, $p< .05$). A main effect of distance was also found (F (1, 122) = 82.10; $p < .0001$; η^2_p: .40), indicating that, across grades, performance on near inferences was more accurate than on far inferences. There was also a main effect of knowledge domain (F (1, 122) = 8.53; $p < .01$; η^2_p: .06), indicating that, across grades, concrete inferences were more accurate than affective inferences. There were no significant interactions (see Table 1.5).

To test if there was a difference in the types of errors children made and whether those errors differed as a function of grade, distance, or knowledge domain a six group (grades 3 through 8) by two error type (processing error versus improbable response) by two-distance (near versus far) by two knowledge domain (concrete versus affective) mixed analysis of variance (ANOVA) was conducted. The analysis revealed a main effect of grade (F (5, 122) =2.36, $p < .05$, η^2_p: .09, paralleling the grade effects for correct responses. There was also a main effect of distance, F (1, 122) = 82.10, $p < .0001$, η^2_p: .40, a main effect of knowledge domain, F (1, 122) = 8.53, $p < .01$, η^2_p: .06, and a main effect of error, F (1, 122) = 487.76, $p < .0001$, η^2_p .80. These main effects were qualified by significant error by distance and error by knowledge domain interactions (F (1, 122) = 81.40, $p < .0001$, η^2_p .40; F (1, 122) = 15.23, $p < .0001$, η^2_p .11, respectively). Paired sample t-tests (Bonferoni correction, $p <.01$) revealed that processing errors (PE) were more common in the far than the near condition, while improbable response errors (IRE) did not differ as a function of distance. Processing errors (PE) were more common in the affective condition than in the concrete condition while the improbable response errors (IRE) did not differ as a function of type of domain knowledge.

Table 1.5

Means (standard deviations) for Accuracy and Errors on Bridging Inferences Task

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
n	19	25	23	21	21	19
Bridge-IT-global	12.79	13.12	13.87	14.76	14.39	15.26
	(2.53)	(2.62)	(3.51)	(2.14)	(2.77)	(3.03)
Bridge-IT-n-c	3.89	4.08	4.00	4.29	4.19	4.37
	(1.29)	(1.00)	(1.31)	(.84)	(.81)	(1.26)
Bridge-IT-n-a	3.63	3.80	3.78	3.95	3.62	4.11
	(.89)	(1.12)	(1.31)	(.92)	(1.36)	(1.10)
Bridge-IT-f-c	2.84	2.84	3.00	3.29	3.24	3.68
	(1.17)	(1.31)	(1.41)	(1.10)	(1.30)	(1.25)
Bridge-IT-f-a	2.42	2.40	3.09	3.24	3.33	3.11
	(1.22)	(1.12)	(1.53)	(.94)	(1.11)	(1.45)
Bridge-IT-nc-pe	.95	.76	.87	.71	.67	.53
	(1.27)	(.93)	(1.18)	(.84)	(.73)	(1.02)
Bridge-IT-na-pe	1.26	1.08	1.09	1.05	1.29	.79
	(.87)	(1.04)	(1.20)	(.92)	(1.27)	(.98)

Bridge-IT-fc-pe	1.95	1.96	1.91	1.71	1.71	1.16
	(1.08)	(1.24)	(1.38)	(1.10)	(1.35)	(1.21)
Bridge-IT-fa-pe	2.47	2.44	1.83	1.71	1.62	1.89
	(1.12)	(1.00)	(1.47)	(1.01)	(1.07)	(1.45)
Bridge-IT-nc-ire	.16	.16	.13	0	.14	.10
	(.37)	(.47)	(.34)		(.36)	(.31)
Bridge-IT-na-ire	.10	.12	.13	0	.09	.10
	(.31)	(.44)	(.34)		(.30)	(.46)
Bridge-IT-fc-ire	.21	.20	.09	0	.05	.16
	(.63)	(.41)	(.29)		(.22)	(.37)
Bridge-IT-fa-ire	.10	.16	.09	.05	.05	0
	(.31)	(.47)	(.42)	(.22)	(.22)	

Note. Bridge-IT-global = raw accuracy scores from Bridging Inferences Task out of 20; Bridge-IT-n-c= raw accuracy scores for near-concrete integration condition out of 5; Bridge-IT-n-a = raw accuracy scores for near-affective integration condition out of 5; Bridge-IT-f-c = raw accuracy scores for far-concrete integration condition out of 5; Bridge-IT-f-a = raw accuracy score for far-affective integration condition out of 5; Bridge-IT-nc-pe = total number of processing errors in near-concrete integration condition out of 5; Bridge-IT-na-pe = total number of processing errors in near-affective integration condition out of 5; Bridge-IT-fc-pe = total number of processing errors in far-concrete integration condition out of 5; Bridge-IT-fa-pe = total number of processing errors in far-affective integration condition out of 5; Bridge-IT-nc-ire = total number of improbable response errors in near-concrete integration condition out of 5; Bridge-IT-na-ire = total number of improbable response errors in near-affective integration condition out of 5; Bridge-IT-fc-ire = total number of improbable response errors in far-concrete integration condition out of 5; Bridge-IT-fa-ire = total number of improbable response errors in far-affective integration condition out of 5.

Relations between Bridging Inferences, Semantic Reactivation, and Semantic Updating

Means (standard deviations) of semantic updating and semantic reactivation are included in Table 1.6. Correlations between the bridging inference outcomes (near-concrete, far-concrete, near-affective, far-affective), semantic reactivation, and semantic updating are reported in Table 1.7. Bridge-IT-near-concrete and Bridge-IT-far-concrete was related to semantic updating while Bridge-IT-near-concrete, -far-concrete, and -far-affective were related to semantic reactivation.

Table 1.6

Means (standard deviations) for Semantic Updating and Semantic Reactivation Tasks

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
n	19	25	23	21	21	19
Reactivation-total	19.68	23.08	23.83	25.76	27.62	29.95
	(5.56)	(6.30)	(6.25)	(6.16)	(8.33)	(4.86)
Updating-total	6.63	7.08	7.30	7.76	8.86	9.42
	(2.98)	(3.60)	(3.18)	(2.88)	(3.85)	(3.34)

*Note.*Reactivation total = total number of words (raw scores) recalled across all conditions; the Updating total = total number of words (raw score) recalled across all conditions.

Table 1.7

Relations of Bridge-IT outcomes, Semantic Reactivation, and Semantic Updating

	NC	NA	FC	FA	SU	SR
NC	--					
NA	.149	--				
FC	.146	.127	--			
FA	.194*	.050	.243**	--		
SU	.194*	.029	.283**	.132	--	
SR	.177*	.076	.342**	.350**	.315**	--

Note. NC = Bridge-IT-near-concrete scores; NA = Bridge-IT-near-affective scores; FC = Bridge-IT-far-concrete scores; FA =

Bridge-IT-far-affective scores; SU = semantic updating; SR = semantic reactivation.

* Correlations are significant at the .05 level (2-tailed test)

** Correlations are significant at the .01 level (2-tailed test)

Working memory component variables (updating and reactivation) were correlated with at least one of the outcome variables, and therefore were included in the hierarchical regressions. In all regression analyses to follow, age was entered first to control for age related effects, as some of the measures were not standardized for age. Raw scores from the TOWRE and the WJIII-voc were entered second to control for word decoding skills and vocabulary knowledge. Raw scores from the semantic updating task were entered in the third block and scores from the semantic reactivation were entered in the fourth and final block. Semantic updating and semantic reactivation were also entered in reverse order, for a total of two regressions for each outcome variable (near-concrete, near-affective, far-concrete, far-affective).

The models predicting Bridge-IT-near-concrete and Bridge-IT-near-affective scores were not significant (F (5, 122) = 1.51) and (F (5, 122) = 1.27), respectively. See Tables 1.A.1 and 1.B.1for regression tables.

The model predicting Bridge-IT-far-concrete scores accounted for 17.6% of the variance (F (5, 122) = 5.21, $p < .001$). The first block included age ($R^2 = .031$, $p < .05$; final $\beta = -.120$, *ns*); followed by word decoding and vocabulary in the second block ($\Delta R^2 = .066$, $p < .05$); semantic updating in the third block ($\Delta R^2 = .046$, $p < .05$; final $\beta = .192$, $p < .05$); and semantic reactivation in the fourth block ($\Delta R^2 = .033$, $p < .05$; final $\beta = .222$, $p < .05$). When semantic reactivation was entered in block three, it accounted for $\Delta R^2 = .048$, $p < .01$; final $\beta = .222$, $p <.05$ and when semantic updating was entered in the last block, it accounted for $\Delta R^2 = .032$, $p < .05$; final $\beta = .192$, $p < .05$. Examination of the final beta coefficients demonstrated that neither word decoding nor vocabulary knowledge were significant predictors, ($t(126) = -.038$, *ns*; final $\beta = -.004$, *ns*) and ($t(126) = 1.861$, *ns*; final $\beta = .220$, *ns*), respectively. See Table 1.C.1 for regression table.

The model predicting Bridge-IT-far-affective scores accounted for 16% of the variance (F (5, 122) = 4.66, $p < .001$). The first block included age ($R^2 = .076$, $p < .001$; final β = .131, *ns*), followed by word decoding and vocabulary knowledge in the second block ($\Delta R^2 = .021$, *ns*), semantic updating in the third block ($\Delta R^2 = .001$, *ns*; final β = -.011, *ns*), and semantic reactivation in the fourth block ($\Delta R^2 = .062$, $p <$ .01; final β = .304, $p < .01$). When semantic reactivation was entered in block three, it accounted for ΔR^2 = .063, $p < .01$; final β = .304, $p < .01$ and when semantic updating was entered in the last block, it accounted for $\Delta R^2 = .000$, *ns*; final β = -.011, *ns*. Examination of the beta coefficients demonstrated that neither word decoding nor vocabulary knowledge were significant predictors, (t(126) = 1.503, *ns*; final β = .169, *ns*) and (t(126) = -1.265, *ns*; final β = -.151, *ns*), respectively. See Table 1.D.1 for regression table.

In summary, semantic reactivation and semantic updating were correlated with near-concrete inferencing but neither of the working memory component measures were related to performance on near inference items in regression models after controlling for age, word decoding and vocabulary knowledge. Semantic updating and semantic reactivation were significant unique predictors of Bridge-IT-far-concrete scores, while semantic reactivation was a significant unique predictor of Bridge-IT-far-affective scores after controlling for age, word decoding, and vocabulary knowledge.

Discussion

The current study demonstrated that the ability to make bridging inferences changes with age and that making bridging inferences across larger chunks of text is more difficult than between adjacent sentences. The findings also supported the hypothesis that making bridging inferences with information related to character motivations, goals, and desires (affective inferences) is more difficult compared to making bridging inferences with concrete causal information (concrete inferences). The working memory processes that predicted performance on the Bridge-IT differed mainly as a function of text distance and, to some extent, as a function of knowledge domain.

As predicted, the present study demonstrated that accuracy scores on the bridging inferences task (Bridge-IT) increased with age. Younger children were less accurate in making bridging inferences compared to older children. This finding is consistent with previous findings in the literature (Barnes et al., 1996; Casteel & Simpson, 1991) and replicates age related effects found with a previous version of the Bridge-IT (Johnston et al., 2008) and an illustrated version of the Bridge-IT (Pike et al., 2010). This finding is unique and extends our understanding of bridging inferences because the Bridge-IT tests within text inferencing and includes variations in text distance (near and far) *and* domain knowledge (concrete and affective).

Effects of Text Distance and Knowledge Domain on Children's Bridging Inference-Making.

Manipulations of text distance affected inference-making accuracy (Barnes et al., 2004; Cain et al., 2003, Johnston et al., 2008). Making bridging inferences across larger chunks of text was more difficult than between contiguous sentences for children across grades and for both knowledge domains. Textual distance effects replicate previous findings in a study using the Bridge-IT, but where knowledge domain was not manipulated (Johnston et al., 2008). In studies of poor comprehenders, increased text distance affected good decoders/poor comprehenders to a greater extent than good decoders/good comprehenders on other types of inference tasks (Barnes et al., 2004; Cain Oakhill, Barnes, & Bryant, 2001; Cain et al., 2003; Cain, Oakhill, & Lemmon, 2004). It appears that typically developing children are equally affected by text distance across knowledge domains in the present study.

Exploring the types of errors children make allows for the indirect examination of the cognitive processes that are believed to support bridging inferences. It was expected that children would make more processing errors (PE) in the Bridge-IT-far condition because it may be more difficult to suppress recently read information that competes with information that is presented earlier, but which is needed to make the inference, and/or it may be more difficult to retrieve information that was read several sentences previously, but which is important for making the correct inference. There were a greater number of processing errors (PE) than improbable answers (IRE) overall; however, processing errors (PE) increased as a function of distance while other errors did not. These findings provide indirect evidence that either suppression or reactivation (or both) are important for making bridging inferences, particularly those that require the integration of information across larger chunks of text (Barnes et al., 2004).

Manipulations of knowledge domain also affected bridging inference making and the types of errors that were made. Inferences based on concrete, causal information (concrete inferences) were more accurate than those based on information about character goals, desires, or motivations (affective inferences). Although this is the first study to examine the effects of knowledge domain on bridging inference-making in children, the findings are consistent with those from studies in which school aged children tended to recall concrete information from narratives more often than internal states of characters (van den Broek, Lorch, & Thurlow, 1996), and from studies of adults in which inferences based on causal knowledge were more easily made than those based on affective knowledge (e.g., Shears & Chiarello, 2004). Additionally, when readers made errors in the current study they tended to choose the answer that was associated with the non-target model more often in the affective than the concrete condition, suggesting that somewhat different processes are engaged when processing affective compared to concrete information. A discussion of why affective inferences might involve a coordination of different cognitive processes than concrete inferences is provided below.

The Relation of Memory Processes to Bridging Inferences.

The correlational analyses revealed that semantic updating was related to Bridge-IT-near-concrete and -far-concrete, while semantic reactivation was related to Bridge-IT-near-concrete and both the concrete and affective Bridge-IT-far conditions. The regression analyses revealed that neither semantic updating nor semantic reactivation was related to near inferences when age, word reading and vocabulary knowledge were controlled in the models. In contrast, semantic reactivation was a unique predictor of Bridge-IT-far inferences, regardless of domain knowledge. These findings suggest that reactivation may be an important process in making inferences across chunks of text, regardless of domain knowledge. To understand these findings consider what the Bridge-IT measures. Each story on the Bridge-IT allows the construction of two competing mental models. In the Bridge-IT-near condition, the reader forms one model in reading sentences one through four and another mental model on sentence five, immediately before being asked to choose the next best sentence in the story which is consistent with the second model. In the Bridge-IT-near condition, therefore, there was no need to reactivate or retrieve information presented earlier in the paragraph in order to make the correct bridging inference. Conversely, in the Bridge-IT-far condition, the reader was presented with and constructed one mental model from sentence one and another mental model in sentences two through five. It is retrieval or reactivation of the first mental model (from sentence one) that is needed to make the correct bridging inference. Therefore, memory reactivation processes are more likely to be needed in the Bridge-IT-far conditions versus the Bridge-IT-near conditions.

The idea that making affective bridging inferences may require the coordination of somewhat different cognitive processes than those required to make concrete bridging inferences was also supported by the regression findings. Semantic updating was a significant predictor of far-concrete inferencing but not far-affective inferencing. When considering the example presented in Table 1.1, it is highly unlikely that a girl with a broken leg would compete in a skipping rope competition. In addition, it is highly unlikely that the condition of a broken leg that occurred last week will change in the near future. Therefore, the process of updating may be important for making concrete inferences because it allows the reader to revise their mental model to include *relevant* information and, if needed, suppress irrelevant information, thereby decreasing the amount of irrelevant information held in working memory. In contrast, updating processes may not be as important for making affective inferences because affective inferences may involve evaluating conflicting information that can be simultaneously true. When considering the example described in Table 1.2, it is possible that a child can be frightened of sharks and large fish in the ocean but also be motivated to face that fear and swim with her friends anyway. Individuals can hold numerous and conflicting desires, goals, or emotions at the same time.

Gernsbacher proposed that the reader can shift attention to build an entirely new model while not entirely suppressing the no-longer-relevant model (Gernsbacher, Robertson, Palladino, & Werner, 2004). It may be that it is in the readers' best interest, in terms of constructing coherent mental models, to hold both models (perhaps one enhanced more than the other) in memory until the situation is resolved (or the item is completed). For this reason, affective items may not require readers to update their mental model per se, but rather to switch attentional focus (Engle, 2002; Unsworth & Engle, 2007) and build a second model until the unfolding context helps to resolve the conflicting information. The interaction of error type and knowledge domain is consistent with this idea. Readers more often erroneously selected the choice that was associated with the non-target mental model for affective than concrete inferences, suggesting that this mental model may not have been entirely suppressed but perhaps strategically maintained in working memory. Affective inference making may therefore require somewhat different cognitive resources compared to concrete inference making (Shears et al., 2007).

Additionally, readers' general knowledge (including personal experience) could affect the content of readers' mental models (Caillies, Denhiere, & Kintsch, 2002; Molinari et al., 2009) and subsequently could influence how readers responded to the dilemmas presented in the motivational items. Perhaps some of the children responding to the situation depicted in Table 1.2 felt that it is more important to face a fear than avoid a fun activity with friends, while others did not. Personal beliefs and experiences may have influenced the process through which children responded to motivational items. Perhaps successfully resolving situations related to characters' internal states in relation to one's own experience requires different cognitive processing than resolving situations that are more concrete (Shears et al., 2007).

Summary, Limitations, and Future Directions

Although the current study revealed interesting findings about bridging inference making in school aged children, the findings must be interpreted with caution for several reasons. First, the length of each item on the Bridge-IT may have influenced how the reader interacted with the text. For example, the passages included in the Bridge-IT were relatively short (five sentences), and thus may not have required the reader to go beyond the text to form a situational model (Graesser, Singer, & Trabasso, 1994) in the same way as would occur when reading longer texts. Reading longer texts may require readers to engage in different strategies to ensure they construct meaning that may rely on different cognitive processes. Future research may focus on investigating the effects of text distance and domain knowledge in longer texts that may better reflect the type of reading situations children and adolescents encounter in real life situations. Additionally, the Bridge-IT was designed to be a conveniently administered, paper-pencil task that is easily scored for accuracy. However, recording response times could provide additional information about participant responses in different conditions related to text distance and knowledge domain (e.g, Gernsbacher, Keysar, Robertson, & Werner, 2001).

The sample size was relatively small given the number of grades tested in the study. Future research should replicate the current findings with a larger number of school-aged children. It would be interesting to explore the relation of other working memory processes such as attentional switching as well as visual working memory to bridging inference making. As discussed above, attentional switching may be an important process when dealing with the type of knowledge included in the Bridge-IT-affective items, where keeping multiple models active until the context helps to resolve the dilemma.

References

Baddeley, A. D. (1986). *Working Memory*. Oxford: Clarendon Press.

Barnes, M. A., Dennis, M., & Haefele-Kalvaitis, J. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inferencing in children from six to fifteen years of age. *Journal of Experimental Child Psychology, 61*(3), 216-241.

Barnes, M. A., Faulkner, H. Wilkinson, M., & Dennis, M. (2004). Meaning construction and integration in children with hydrocephalus. *Brain and Language, 89*, 47 - 56.

Barnes, M. A., Huber, J., Johnston, A. M., & Dennis, M. (2007). A model of comprehension in spina bifida meningomyelocele: Meaning activation, integration, and revision. *Journal of the International Neuropsychological Society, 13*, 854 – 864.

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

Cain, K., Oakhill, J. V., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language, 30*, 681 - 694.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96*, 671–681.

Caillies, S., Denhiere, G., & Kintsch, W. (2002). The effect of prior knowledge on understanding from text: Evidence from primed recognition. *European Journal of Cognitive Psychology, 14*(2), 267 - 286.

Campion, N., & Rossi, J. P. (2001). Associative and causal constraints in the process of generating predictive inferences. *Discourse Processes, 31*, 263–291.

Carretti, B., Cornoldi, C., De Beni, R., & Romano, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology, 91*, 45 - 66.

Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J., Willcutt, E., & Olson, R. K. (2012, February 20). Predicting Word Reading and Comprehension With Executive Function and Speed Measures Across Development: A Latent Variable Analysis. *Journal of Experimental Psychology: General*, *141*(3), 470 – 488.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19-23.

Field, A. (2005). *Discovering statistics using SPSS.*(2nd ed.). London, UK: Sage.

Garson, N., Bryson, S. E., & Smith, E. M. (2006). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31 – 60.

Gernsbacher, M. A. (1990).*Language comprehension as structure building.* Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Gernsbacher, A. M. (1997). Attentuating interference during comprehension: The role of suppression. *The Psychology of Learning and Motivation, 37*, 85 - 104.

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(2), 245 - 262.

Gernsbacher, M. A., Keysar, B., Robertson, R. R. W., & Werner, N. K. (2001). The role of suppression and enhancement in understanding metaphors. *Journal of Memory and Language, 45*, 433 - 450.

Gernsbacher, M. A., Robertson, R. R., Palladino, P., & Werner, N. K. (2004). Managing mental representations during narrative comprehension. *Discourse Processes, 37*(2), 145 – 164.

Gernsbacher, M. A., & Shlesinger, M. (1997). The proposed role of suppression in simultaneous interpretation. *Interpreting, 2*(1/2), 119 - 140.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101* (3) - 371 - 395.

Johnston, A. M., Barnes, M. A., & Davis, J. B. (2007, July). *The Role of Attentional Control in Reading Comprehension.* Presented at the 14th Annual meeting for the Society for the Scientific Study of Reading (SSSR), Prague, Czech Republic.

Johnston, A. M., Barnes, M. A., & Desrochers, A. (2008, June). Bridging Inferences in School Aged Children. Poster presented at the 39th Annual Convention of the Canadian Psychology Association (CPA), Halifax, NS.

Kintsch, W., & Kintsch, E. (2005). Comprehension. In S. G. Paris and S. A. Stahl (Eds.), *Children's Reading Comprehension and Assessment* (pp. 71 - 92). New York: Lawrence Erlbaum.

Linderholm, T., Gernsbacher, M. A., van den Broek, P., Neninde, L., Robertson, R. R. W., & Sundermier, B. (2004). Suppression of story character goals during reading. *Discourse Processes, 37*(1), 67 - 78.

Lynch, J. S., & van den Broek, P. (2007). Understanding the glue of narrative structure: Children's on- and off-line inferences about characters' goals. *Cognitive Development, 22*, 323 - 340.

Long, D. L., Oppy, B. J., &Seely, M. R. (1997). Individual differences in readers' sentence- and text-level representations. *Journal of Memory and Language, 36*, 129 – 145.

Molinari, C., Burin, B., Saux, G., Barreyro, J. P., Irrazabal, N., Bechis, M. S., Duarte, A., & Ramenzoni, V. (2009). Fictional characters' emotional state representation: What is its degree of specificity? *Psicothema, 12*(1), 9 – 14.

O'Brien, E. J., Cook, A. E., & Gueraud, S. (2010). Accessibility of outdated information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(4), 979 – 991.

O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1200 - 1210.

Palladino, P., Cornoldi, C., DeBeni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory & Cognition, 29*(2), 344-354.

Pike, M. M., Barnes, M. A., & Barron, R. W. (2010). The role of illustrations in children's inferential comprehension. *Journal of Experimental child Psychology, 105*, 342 – 255.

Pimperton, H., & Nation, K. (2010). Suppressing irrelevant information from working memory: Evidence for domain-specific deficits in poor comprehenders. *Journal of Memory and Language, 62,* 380 – 391.

Rapp, D. N., & van den Broek, P. (2005).Dynamic text comprehension. *Current Directions in Psychology Science, 14*(5), 276 - 279.

Shears, C., & Chiarello, C. (2004). Knowledge-based inferences are NOT general. *Discourse Processes, 38*, 31–55.

Shears, C., Hawkins, A., Varner, A., Lewis, L., Heatley, J., & Twachtmann, L. (2008). Knowledge-based inferences across the hemispheres: Domain makes a difference. *Neuropsychologia, 46*(10), 2563 – 2568.

Shears, C., Miller, V., Ball, M., Hawkins, A., Griggs, J., & Varner, A. (2007). Cognitive demand differences in causal inferences: Characters' plans are more difficult to comprehend than physical causation. *Discourse Processes, 43*(3), 255 – 278.

Singer, M., Halldorson, F. C., Lear, C., & Andrusiak, P. (1992). Validation of Causal Bridging Inferences in Discourse Understanding. *Journal of Memory and Language, 31*, 507 - 524.

Torgensen, J. K., Wagner, R. K. & Rashotte, C. A. (1999).*Test of Word Reading Efficiency* (*TOWRE*).

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*(1), 104 – 132.

van den Broek, P. (1994). Comprehension and memory of narrative texts: Inferences and coherence. In M. A. Gernsbacher (Ed.). *Handbook of psycholinguistics* (pp. 539-588). San Diego, CA, US: Academic Press.

van den Broek, P., Kendeou, P., Kremer, K., Lynch, J. S., Butler, J., White, M. J., & Lorch, E. P. (2005). Assessment of comprehension abilities in young children. In S. Stahl & S. Paris (eds.), *Children's Reading Comprehension and Assessment*, (pp.107-130). Mahwah, NJ: Erlbaum.

van den Broek, P., Lorch, E. P., & Thurlow, R. (1996). Children's and adults' memory for television stories: The role of causal factors, story-grammar categories, and hierarchical level. *Child Development, 67*, 3010–3028.

Woodcock, R., Mather, N., & McGrew (2001).*Woodcock-Johnson III – Test of Achievement.* Chicago: Riverside.

Appendix 1.A

Table 1.A.1

Summary of Hierarchical Regression Predicting Bridging Inference Task Near-concrete condition (*N* = 128)

Outcome	Step	Variable	ΔR^2	B	SE B	β
Near-concrete						
	1	Age	.021	.015	.078	.024
	2	Word Decoding	.010	-.002	.012	-.024
		Vocabulary		.027	.040	.085
	3	Updating	.022	.046	.030	.144
	4	Reactivation	.005	.013	.017	.085
	3	Reactivation	.009	.013	.017	.085
	4	Updating	.018	.046	.030	.144

Appendix 1.B

Table 1.B.1

Summary of Hierarchical Regression Predicting Bridging Inference Task Near-affective condition (*N* = 128)

Outcome	Step	Variable	ΔR^2	B	SE B	β
Near-affective						
	1	Age	.005	-.072	.081	-.113
	2	Word Decoding	.044	.026	.013	.249*
		Vocabulary		.018	.042	.056
	3	Updating	.001	-.009	.031	-.026
	4	Reactivation	.000	.001	.017	.005
	3	Reactivation	.000	.001	.017	.005
	4	Updating	.001	-.009	.031	-.026

$p^* < 0.05$

Appendix 1.C

Table 1.C.1

Summary of Hierarchical Regression Predicting Bridging Inference Task Far-concrete condition (*N* = 128)

Outcome	Step	Variable	ΔR^2	B	SE B	β
Far-concrete						
	1	Age	.031*	-.086	.085	-.120
	2	Word Decoding	.066*	.000	.013	-.004
		Vocabulary		.082	.044	.220
	3	Updating	.046*	.072	.033	.192*
	4	Reactivation	.033*	.040	.018	.222*
	3	Reactivation	.048**	.040	.018	.222*
	4	Updating	.032*	.072	.033	.192*

$p^* < 0.05$; $p^{**} < 0.01$

Appendix 1.D.1

Table 1.D.1

Summary of Hierarchical Regression Predicting Bridging Inference Task Far-affective condition (N = 128)

Outcome	Step	Variable	ΔR^2	B	SE B	β
Far-affective						
	1	Age	.076**	.094	.086	.131
	2	Word Decoding	.021	.020	.013	.169
		Vocabulary		-.056	.044	-.151
	3	Updating	.001	-.004	.033	-.011
	4	Reactivation	.062**	.055	.019	.304**
	3	Reactivation	.063**	.055	.019	.304**
	4	Updating	.000	-.004	.033	-.011

$p^* < 0.05$; $p^{**} < 0.01$

Chapter 3: The Relation of Working Memory Processes and Bridging Inferences to Reading Comprehension in Children in Grades Three through Eight

Reading comprehension is a dynamic process in which readers interact with text to form coherent mental representations of the situation described by the text (Kintsch & Kintsch, 2005; Snow, 2002). Historically, different reading comprehension measures have been used interchangeably to measure and draw conclusions about the complex process of reading comprehension (Keenan, Betjemann, & Olson, 2008). However, the skills and abilities needed to demonstrate adequate comprehension skill may vary depending on the measure being used to assess reading comprehension (e.g., Bowyer-Crane & Snowling, 2005; Cutting & Scarborough, 2006; Francis, Fletcher, Catts, & Tomblin, 2005; Francis, Snow, August, Carlson, & Iglesias, 2006; Johnston, 2006; Keenan et al., 2008; Nation & Snowling, 1997).

Standardized measures are used to assess ability and identify deficits that result in children's failure to understand what they read (Francis et al., 2006). Cutting and Scarborough (2006) noted in one study (Rimrodt, Lightman, Roberts, Denckla, and Cutting, 2005) that although almost half of their sample performed below their age expected range on one of three reading comprehension measures, less than 10% were identified as having poor comprehension by all three measures. Such findings highlight a potential problem in psychoeducational assessment: difficulties in reading comprehension might be overlooked if the reading comprehension measure used does not provide adequate coverage of the various skills needed to understand text.

Nation and Snowling (1997) were among the first to investigate the skills and processes that support performance on different reading comprehension measures. As researchers began to explore which skills were most related to reading comprehension performance (e.g., Francis et al., 2005; Keenan et al., 2008; Nation & Snowling, 1997), they consistently replicated the well-established finding that word-decoding skills are important for comprehension. However, the importance of decoding varies depending on the measure. Word decoding skills are more strongly related to measures with a cloze format (sentence completion) whereas decoding *and* listening comprehension are related to measures with multiple-choice format (Francis et al., 2005; Keenan et al. 2008; Nation & Snowling, 1997). Keenan et al. (2008) suggested that it is not simply response format that accounts for differences between measures but instead it is the length of the passage that influences which skills best support performance on different comprehension tasks. The cloze and multiple-choice tasks used by Keenan et al. (2008) were short (one to two sentences) whereas the measures used by Nation and colleagues (1997) contained longer passages. Keenan et al. (2008) suggested that shorter passages are more reliant on decoding skills because there is less information presented in the text to support gap-filling, such as using context to infer the meaning of a word when one is unable to decode it. Others have discussed the possibility that different comprehension tests measure different aspects of the complex process of reading comprehension (van den Broek, Rapp, & Keough, 2005).

Reading Comprehension: Word- and Text-Level Skills, and Working Memory

Reading comprehension relies upon skilled word decoding and listening comprehension (Gough & Tunmer, 1986; Hulme & Snowling, 2011; Kirby & Savage, 2008). Accurate decoding is essential for comprehension because if the reader cannot decode, the semantics of the text (e.g., word meanings, syntactic knowledge) cannot be accessed from long-term memory. Being able to read words fluently is also important because if decoding is slow and labored, the reader must hold information longer in memory before integration can begin, tying up cognitive resources that could be allocated to other processes that are important for comprehension (Wolf & Katzir-Cohen, 2001).

When decoding skills are well developed, comprehension difficulties can still manifest (Barnes, Huber, Johnston, & Dennis, 2007; Oakhill, 1993). Cognitive models of reading comprehension (e.g., Gernsbacher, 1990; Kintsch, 2002; van den Broek et al., 2005) describe how the process of comprehension occurs as the text unfolds and how comprehension is supported by various skills and cognitive processes such as inference and working memory. In terms of individual differences in comprehension, some suggest that poor comprehension skills are mostly due to underlying weaknesses in basic oral language skills such as vocabulary knowledge, grammatical understanding, and language comprehension (Hulme & Snowling, 2011; Nation & Snowling, 2000), while others suggest a role for text- or discourse-level processes such as inference making, and general cognitive abilities such as working memory (Cain, Oakhill, & Lemmon, 2004; Carretti, Borella, Cornoldi, & DeBeni, 2009; Christopher, Miyake, Keenan, Pennington, DeFries, Wadsworth, Willcutt & Olson, 2012; Schmalhofer, McDaniel, & Keefe, 2002). The literature on relations of working memory and inference to reading comprehension is reviewed in more detail below.

Working Memory and Comprehension.

Baddeley (1986) proposed that working memory is the cognitive system that provides a mental workspace in which information-processing tasks are executed, such as comprehending written text. Working memory is a strong predictor of children's reading comprehension performance and those with comprehension difficulties have been found to perform more poorly on working memory tasks (Cain et al., 2004). Researchers have suggested that the relation between reading comprehension and working memory is a reflection of the relation of *specific components* of working memory with comprehension (Christopher et al., 2012). Christopher and colleagues investigated the unique relation of working memory and working memory components that consisted of suppression/inhibition, processing speed, and word naming speed with reading comprehension and word reading outcome measures. Their findings demonstrated that both working memory and processing speed predicted unique variance in both word reading and reading comprehension, highlighting the importance of being able to hold and manipulate information in memory and quickly process visual information for both word reading and reading comprehension. Because working memory has a limited capacity to store and manipulate the large amount of information that becomes activated during reading, it becomes increasingly important that working memory resources are efficiently allocated during reading.

When reading begins, each accurately decoded word activates knowledge from long-term memory, such as word meanings and other general knowledge associated with the text (Gernsbacher & Faust, 1991; Kintsch, 2002; van den Broek, Young, Tzeng, & Linderholm, 1999; van den Broek et al., 2005). Activated information may be either contextually relevant or irrelevant and the relevance of some of this activated information may change over the course of the text; as such, context is thought to be pivotal in guiding the maintenance of information in working memory (Gernsbacher, 1990). Irrelevant information is suppressed or pushed out of the focus of attention (Gernsbacher, 1990; Linderholm, Gernsbacher, van den Broek, Neninde, Robertson, & Sundermier, 2004) to prevent it from being integrated into the reader's mental model and hindering comprehension. The content relevant information held in memory is integrated to form a mental model of the situation described by the text.

Meaning construction is an ongoing process that occurs in reading cycles (Kintsch & Kintsch, 2005; Linderholm, Virtue, Tzeng, & van den Broek, 2004). Activated information and mental models are held in working memory as they are carried from one reading cycle to the next. The mental model is updated or revised to include newly processed information that supports comprehension (Kintsch, 1988, 2002; Perfetti, Yang, & Schmalhofer, 2008; Zwaan & Madden, 2004). If the mental model is not updated to include important information, the reader will not gain a full appreciation for what is being communicated and comprehension may begin to break down (van den Broek et al., 2005).

The role of updating and reactivation processes in comprehension. The process of *updating* information is related to reading comprehension (Palladino, Cornoldi, DeBeni, & Pazzaglia, 2001) and findings support the hypothesis that *updating* is a specific aspect of working memory that mediates the relation between working memory and reading comprehension (Carretti, Cornoldi, De Beni, & Romano, 2005). Palladino et al., (2001) investigated the process of updating with a task in which participants were asked to recall the three smallest objects from a list of words that was read to them. To be successful, participants retained an updated list by adding and dropping words according to the relative size of their real-life referents. Palladino et al. (2001) found that participants who were less skilled in reading comprehension recalled fewer objects in the correct order on the updating task than their better comprehending peers.

As the text unfolds and the reader continues to construct meaning, new ideas are often presented that are highly related to previously processed ideas, are no longer active in working memory (Kintsch, 2002). Words or sentences act as cues to *reactivate* (or bring back into working memory) information that was processed during previous reading cycles (Albrecht & O'Brien, 1993; Gernsbacher, 1990; Kintsch, 1991; Linderholm et al., 2004). These cues can include explicitly restating words or paraphrasing an attribute, action, or event that is associated with a character that was previously introduced (O'Brien, Rizzella, Albrecht, & Halleran, 1998; Perfetti et al., 2008). The more highly associated the previously encountered information is to the most recently encoded text, the faster the reader is able to reactivate or retrieve it (Gernsbacher, Robertson, Palladino, & Werner, 2004). Much research has been conducted regarding the availability or ease of retrieval of processed information as one continues to read (e.g., Cook, Gueraud, Was, & O'Brien, 2007; Linderholm et al. 2004).

Working memory is related to reading comprehension, and components of working memory such as updating and reactivation have been individually related to reading comprehension or hypothesized to be related to comprehension and specific text-level processes such as inference. However, no study has investigated the specific roles of both updating and reactivation in relation to a number of broad reading comprehension outcomes.

The role of bridging inferences in reading comprehension. Inferencing involves integrating information within text as well as between text and general knowledge. Inferencing consistently accounts for unique variance in reading comprehension in skilled adult readers (Schmalhofer et al., 2002), and in typically developing children (Pike, Barnes, & Barron, 2010), and differentiates children who are characterized with poor comprehension and adequate word decoding from those with good comprehension (Barnes, Johnston, & Dennis, 2007; Barnes, Faulkner, Wilkinson, & Dennis, 2004; Cain & Oakhill, 1999; Cain, Oakhill, & Bryant, 2004; Cain, Oakhill, & Elbro, 2003; Cain et al., 2004; Long et al., 1997; Oakhill, 1993).

Although previous research has explored relations between word decoding and listening comprehension to different measures of reading comprehension (e.g., Keenan et al., 2008), little research has explored the relations of working memory and inference-making in addition to word-level skills and knowledge (i.e., word decoding, vocabulary knowledge) to different reading comprehension measures. The current study explores the contributions of verbal working memory span, two specific working memory processes, semantic updating and semantic reactivation, and bridging inference making to performance on three reading comprehension outcomes - Passage Comprehension subtest from the Woodcock-Johnson III Test of Achievement, WJIII-PC (Woodcock & Johnson, 2001), Paragraph Reading subtest from the Test of Reading Comprehension, third edition (TORC3-PR, Brown, Hammill, & Wiederholt, 1995), and Reading Comprehension subtest from the Wechsler Individual Achievement Test, second edition (WIAT-II-RC, Wechsler, 2002), in children in grades three through eight. Although some of these general purpose cognitive processes (working memory) and text-level skills (inference) have been considered in relation to some reading comprehension outcomes, no study has considered all of these purported key comprehension-related skills together for the prediction of reading comprehension outcomes, nor has any study investigated these skills in relation to the broad range of comprehension measures used here.

Three reading comprehension measures were chosen for the current study for three reasons. First, each measure has a different response format (i.e., cloze task, multiple choice, open ended) and second, the passages vary in length. Both factors 1 and 2 have been argued to underlie the differences found between reading comprehension measures (Keenan et al., 2008; Nation & Snowling, 1997). And third, Johnston (2006) has argued that different RC measures may assess different aspects of mental model building. For example, success on the WJIII-PC (cloze task, generate a missing word to complete a sentence or short passage) might involve building a mental model of the situation described in the text. To complete items on the WJIII-PC, one must hold a partially constructed mental model in memory while retrieving a word to correctly complete the sentence or short passage. In contrast, the TORC3-PR might involve building a mental model and then *reflecting* upon the model when presented with multiple-choice questions at the end of each passage. Success on the TORC3-PR might rely much less on memory for the text because the text is available to the reader, allowing the reader to refer back and search for the correct answer to the question. The reader may not choose to refer back to the text, but instead may refer to the mental representation held in memory while responding to questions. The WIAT-II-RC was included in the current study because, like the TORC3-PR, it has longer passages (greater than 10 sentences) and the text is made available to the reader while responding to questions. Unlike the TORC3-PR, the reader is asked to orally respond to open-ended questions, responses can be queried, and some passages are accompanied by illustrations.

Aims and Hypotheses.

The first aim of the current study was to investigate the relations of working memory and two specific working memory processes (semantic updating and semantic reactivation) to performance on three different standardized reading comprehension measures (i.e., WJIII-PC; TORC3-PR; WIAT-II-RC) in children ranging from grades three through eight while controlling for word-level decoding and vocabulary skills. The second aim was to test whether bridging inferences have additional predictive power for reading comprehension once both word-level and working memory processes have been taken into consideration. Consistent with aim number one, it was hypothesized that semantic updating would uniquely predict reading comprehension scores as measured by the WJIII-PC. Updating was hypothesized to be important for success on the WJIII-PC (cloze task) because it may require the reader to build a mental model (which requires updating) in order to generate word(s) to complete the passage. Second, it was hypothesized that semantic reactivation would uniquely predict scores on the TORC3-PR. Reactivation should be important because, when responding to multiple-choice questions, the reader may reactivate or retrieve particular portions of the constructed mental model to correctly respond to the question. Third, it was hypothesized that scores on semantic reactivation would also uniquely predict scores on the reading comprehension subtest from the WIAT-II-RC. Reactivation was hypothesized to be important on this measure because the WIAT-II-RC requires the reader to generate responses to open ended questions about passages. To answer the questions, the child may reactivate or retrieve particular portions of the text that are no longer in the focus of attention.

Because general measures of working memory have been related to reading comprehension outcomes across many studies, it was considered important to also use a standard measure of working memory span in the current study. Although verbal working memory span is related to reading comprehension outcomes, these measures likely tap many processes including updating, reactivation, suppression and so forth. The complex nature of typical working memory tasks makes it difficult to determine whether some working memory processes are more important than others for reading comprehension. In order to explore the role of verbal working memory span in comprehension, it was hypothesized that performance on a measure of verbal working memory span would uniquely predict performance on all three reading comprehension measures (Gernsbacher & Faust, 1991; Pimperton & Nation, 2010) when semantic updating and semantic reactivation were not in the model. In addition, updating and reactivation, processes hypothesized in the current study to be related to comprehension based on cognitive models of the reading comprehension process, were tested to determine if they account for unique variance in the prediction of reading comprehension with verbal working memory span controlled. Finally, verbal working memory span was tested to determine if it uniquely predicted reading comprehension after accounting for updating and reactivation processes because this knowledge is relevant for understanding the nature of the relation of working memory and reading comprehension.

Consistent with aim number two, it was hypothesized that bridging inferences tap a skill that is important for reading comprehension but cannot be fully accounted for by other comprehension-related abilities such as vocabulary knowledge and working memory. It was predicted that bridging inference scores would predict unique, significant variance on all three reading comprehension measures after word-level, text-level, and working memory skills are controlled.

Method

Participants

One hundred and thirty-eight English-speaking children (70 boys and 68 girls) ranging from grades three through eight (8- to 13-years-old), were recruited from local school boards in south-central Ontario to participate in the present study. The sample used in this study was a subsample of those used in Study 1. In addition to the exclusion criteria used in Study 1, an additional five students who were 14 years old were excluded from the analyses because the standardized working memory task was designed and normed for children 13 years and younger. The final sample consisted of 123 children; there were 19 grade three students (*M* age = 8.21); 25 grade four students (*M* age = 9.16); 23 grade five students (*M* age = 10.39); 21 grade six students (*M* age = 11.38); 21 grade seven students (*M* age = 12.38); and 14 grade eight students (*M* age = 12.93).

Measures

Because several of the measures used in Study 2 are the same as those in Study 1, only measures unique to Study 2 are discussed in detail.

Sight Word Efficiency subtest from the Test of Word Reading Efficiency (TOWRE; Torgensen et al., 1999). The TOWRE was used to obtain a measure of word reading efficiency, which involved accuracy and fluency. Although standardized scores were used in the screening process, raw scores were used in analyses because experimental tasks measuring working memory components (e.g., semantic updating, semantic reactivation) were also used in the analyses. American test norms were used to derive standardized scores.

The Picture Vocabulary subtest from the Woodcock-Johnson III (WJIII-voc; Woodcock et al., 2001). The WJIII-voc was used to obtain a measure of vocabulary knowledge. Each child was presented with a page including six pictures and was asked to name each picture using one word. Raw scores were used in analyses for the same reasons mentioned above. American test norms were used to derive standardized scores.

Bridging Inferences Task (Bridge-IT). The Bridge-IT measures the ability to integrate and make inferences using information presented within text and is described in detail in Chapter 2. For this study the sum of all correct answers across the entire measure, combining near, far, concrete, and affective items, created the Bridge-IT-global score (total correct out of 20), which was used in the regression analyses (see Table 2.5).

The Verbal Working Memory subscale from the Wide Range Assessment of Memory and Learning (WRAML; Sheslow & Adams, 2003). The WRAML measures verbal working memory. Each child completed two levels of items. Each child was presented with a list of animal words and object words. Each child was asked to recall all animals in any order followed by all objects in any order (level A) and all animals first in order of size (smallest to largest) followed by all objects in any order (level B). Raw scores were used. The WRAML has a test-retest reliability of $r = .77$ when tested within 49 days.

The Semantic Reactivation Task. This task was designed to measure the ability to reactivate or retrieve previously processed information when cued to do so. Because the amount of proactive interference (or lack of) encountered while reading may vary depending on length and response format, the total raw score was used in analyses in an attempt to capture the reactivation processes that may be engaged while reading passages that reflect these differences across the three reading comprehension measures. This task is described in greater detail in Chapter 2.

The Semantic Updating Task. This task was designed to measure the ability to update information according to semantic content (i.e., ordering according to size), an important working memory process (Garson, Bryson, & Smith, 2008; Palladino et al., 2001). Because the number of updates made while forming mental models while reading may vary depending on the demands of the text being read, such as length of the text and the response format of the task, the total raw score was used in all analyses in an attempt to capture the updating processes that may be required across the three reading comprehension measures included in this study. The task is described in greater detail in Chapter 2.

Standardized Reading Comprehension Tests

Passage Comprehension subtest from the Woodcock-Johnson III (WJIII-PC; Woodcock et al., 2001). The WJIII-PC was used to obtain a measure of reading comprehension achievement. This commonly used test requires each child to read sentences and passages of increasing length and then generate a missing word to make the passage complete and coherent. As the child progresses, items become more difficult as the passages increase in length, illustrations are removed, vocabulary becomes more sophisticated, syntax becomes more complicated, and semantic clues decrease. Standardized scores (x = 100; SD = 15) were derived from age based on American norms. This test includes norms ranging from 2 years to over 90 years of age and test-retest reliability was $r = .80$ for children between the ages of 2 to 18 years when tested one to two years later. Raw scores were used in the analyses.

Paragraph Reading subtest from the Test of Reading Comprehension, third edition (TORC3-PR; Brown, Hammill, & Wiederhold, 1995). The TORC3-PR was used to obtain a measure of reading comprehension achievement. Each child reads a paragraph and then answers five multiple-choice questions relating to the paragraph while the text is made available. Of the five questions, one assesses the child's understanding of the overall theme of the passage by asking him or her to select the best title. Two questions assess the child's ability to recognize details from the passage. And two questions assess the child's ability to make two types of inferences. One involves making a negative inference, that is, to decide which choice could not go with the story and the other is a knowledge-based inference. Age based American norms ranging from 7 to 17 years, 11 month of age are provided. The manual reports test-retest reliability of $r = .81$ two months following the first administration. Raw scores were used in the analyses.

Passage Comprehension subtest from the Wechsler Individual Achievement Test, second edition (WIAT-II-RC; Wechsler, 2001). The WIAT-II-RC was used to obtain a measure of reading comprehension achievement. Each child is presented with a passage (10 or more sentences) accompanied by illustrations to read and is then asked to verbally respond to open ended questions related to the passage while the passage is made available. Passages become increasing longer; include more advanced vocabulary, and increasingly sophisticated syntax and grammar. Interspersed among longer passages are shorter passages (e.g., one to three sentences) that are not accompanied by illustrations. Children are asked to answer open-ended questions about shorter passages. Questions assess the child's ability to recall stated details from the passage, draw conclusions, make inferences (bridging, knowledge-based, predictive), identify the main idea, sequence events in the text, and recognize facts versus opinions.

Age based norms ranging from 5 to 19 years, 11 months of age are provided. Test-retest reliability was r = .93 when tested, on average, 10 days following the first administration. Raw weighted scores provided in the manual were used for all analyses because children were required to respond to different passages depending on their grade level. Canadian test norms were used for deriving raw weighted scores.

Procedure

Children who passed the word reading screening measure were tested individually on measures of vocabulary knowledge, verbal working memory, semantic reactivation and updating, and tested in small groups on the measure of bridging inferences, in 45-60 minute blocks over the course of several sessions.

Results

Because the participants of Studies 1 and 2 do not completely overlap, characteristics and scores on standardized measures of vocabulary knowledge, word decoding skills, and verbal working memory are presented in Table 2.4. To test if there was a difference in word reading skill, vocabulary knowledge, and verbal working memory between grades, a one-way ANOVA was conducted on each of the standardized measures using standardized scores. Main effects of grade were found for word reading (F (5, 117) = 2.99; $p < 0.05$) and vocabulary knowledge (F (5, 117) = 2.74; $p < 0.05$) but not verbal working memory (F (5, 117) = 1.65; *ns*). Tukey post-hoc tests ($p < .05$) revealed that the grade three word reading standard scores were significantly higher than the standard scores of grade five and the grade three vocabulary knowledge standard scores were significantly higher than the standard scores of grade seven. Means and standard deviations for the Semantic Updating Task, Semantic Reactivation Task, and the Bridging Inference Task are presented in Table 2.5. Means and standard deviations for the standardized measures of reading comprehension are presented in Table 2.6.

Table 2.4

Means (standard deviations) for standardized measures of vocabulary, word reading, and verbal working memory

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
n	19	25	23	21	21	14
TOWRE-raw	65.74	68.36	70.52	78.33	81.71	86.86
	(5.55)	(8.44)	(9.04)	(8.95)	(8.73)	(7.62)
TOWRE-ss	112.79	108.56	101.43	105.86	105.05	108.00
	(7.43)	(10.10)	(10.32)	(11.13)	(10.87)	(10.06)
WJIII-voc-raw	24.68	25.32	26.13	28.43	28.86	31.64
	(1.86)	(2.01)	(2.75)	(2.82)	(2.52)	(3.25)
WJIII-voc-ss	110.37	105.76	103.09	104.81	102.86	109.07
	(7.85)	(7.19)	(7.85)	(9.07)	(7.42)	(10.25)
WRAML-raw	19.74	23.64	26.04	27.43	27.95	28.50
	(5.14)	(3.73)	(4.16)	(5.01)	(4.66)	(4.93)
WRAML-ss	9.42	10.76	11.22	10.95	10.48	10.86
	(2.34)	(1.53)	(1.91)	(2.54)	(2.29)	(2.66)

Note. TOWRE-raw = raw scores from the Sight Word Efficiency subtest of the Test of Word Reading Efficiency (TOWRE); TOWRE-ss = standard scores from the Sight Word Efficiency subtest of the TOWRE; WJIII-voc-raw = raw scores from the Picture Vocabulary subtest of the Woodcock-Johnson test of Achievement (WJIII); WJIII-voc-ss = standard scores from the Picture Vocabulary subtest of the WJIII; WRAML-raw = raw scores from the verbal working memory subtest of the Wide Range Assessment of Memory and Learning; WRAML-ss = standardized scores from the verbal working memory subtest of the Wide Range Assessment of Memory and Learning.

Table 2.5

Means (standard deviations) for Semantic Updating Task, Semantic Reactivation Task, and the Bridging Inferences Task

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
n	19	25	23	21	21	14
Reactivation-	19.68	23.08	23.83	25.76	27.62	29.57
global	(5.56)	(6.30)	(6.25)	(6.16)	(8.33)	(5.30)
Updating-	6.63	7.08	7.30	7.76	8.86	8.43
global	(2.98)	(3.60)	(3.18)	(2.88)	(3.85)	(2.47)
Bridge-IT-	12.79	13.12	13.87	14.76	14.39	14.86
global	(2.53)	(2.62)	(3.51)	(2.14)	(2.77)	(2.98)

Note. Reactivation-global = total number of words (raw scores) recalled across the Semantic Reactivation Task; Updating-global = total number of words (raw score) recalled across the Semantic Updating Task; Bridge-IT-global = accuracy scores from the Bridging Inference Task out of 20.

Table 2.6

Means (standard deviations) for Standardized Reading Comprehension Tasks

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
n	19	25	23	21	21	14
WJIII-PC-raw	26.79	27.88	30.39	30.62	32.29	33.36
	(2.59)	(2.67)	(3.23)	(2.33)	(2.03)	(3.03)
WJIII-PC-ss	101.42	97.28	98.17	94.57	96.05	97.14
	(6.18)	(6.05)	(8.42)	(6.05)	(5.50)	(8.24)

TORC3-PR-raw	10.95	11.56	13.65	17.14	15.57	20.29
	(3.63)	(2.99)	(6.61)	(6.89)	(6.30)	(7.47)
TORC3-PR-ss	11.37	10.52	9.35	9.71	9.38	9.93
	(1.64)	(1.16)	(2.10)	(2.05)	(2.13)	(2.27)
WIAT-II-RC-wr	126.37	124.96	135.65	142.00	156.95	154.86
	(16.43)	(11.20)	(14.68)	(23.60)	(23.98)	(34.22)
WIAT-II-RC-ss	110.79	103.79	106.26	103.71	107.00	99.57
	(10.24)	(10.47)	(11.43)	(17.07)	(11.26)	(16.55)

Note. WJIII-PC-raw = raw scores from the Passage Comprehension subtest from the WJIII-PC; WJIII-PC-ss = standardized scores from the Passage Comprehension subtest from the WJIII-PC; TORC3-PR-raw = raw scores from the Paragraph Reading subtest from the TORC3-PR; TORC3-PR-ss = standardized scores from the Passage Reading subtest from the TORC3-PR; WIAT-II-RC-wr = weighted raw scores from the Reading Comprehension subtest from the WIAT-II-RC; WIAT-II-RC-ss = standardized scores from the Reading Comprehension subtest from the WIAT-II-RC.

Relations of Working Memory, Semantic Reactivation, Semantic Updating, and Bridging Inferences to Reading Comprehension

Correlations between the control variables (age, TOWRE, WJIII-voc), reading comprehension measures (WJIII-PC, TORC3-PR, WIAT-II-RC), semantic updating, semantic reactivation, verbal working memory, and Bridge-IT total accuracy scores are reported in Table 2.7. Age, word reading (TOWRE), vocabulary knowledge (WJIII-voc), verbal working memory (WRAML), semantic updating, semantic reactivation, and Bridge-IT scores were significantly correlated with all three standardized reading comprehension measures (WJIII-PC, TORC3-PR, WIAT-II-RC), with the exception of semantic updating, which was not correlated with the TORC3-PR. Partial correlations, controlling for age, revealed significant correlations between word decoding (TOWRE), vocabulary knowledge (WJIII-voc), verbal working memory (WRAML), semantic reactivation, and Bridge-IT scores with all three reading comprehension measures (WJIII-PC, TORC3-PR, WIAT-II-RC). Semantic updating was correlated with one of the comprehension measures (WJIII-PC).

Table 2.7

Correlations between reading comprehension outcomes, working memory processes, and bridging inferences

	Age	TOWRE	WJIII-voc	WJIII-PC	TORC 3-PR	WIAT-II-RC	Updat-ing	React-ivation	WRAML	Bridge-IT-global
Age	--									
TOWRE	.627**	--	.345**	.271**	.237**	.320**	.186*	.210*	.009	.219*
WJIII-voc	.595**	.579**	--	.492**	.253**	.441**	.092	.349**	.232*	.212*
WJIII-PC	.622**	.559**	.679**	--	.257**	.497**	.279**	.380**	.259**	.462**
TORC3-PR	.444**	.430**	.446**	.456**	--	.346**	.072	.280**	.210*	.339**
WIAT-II-RC	.505**	.505**	.606**	.650**	.492**	--	.165	.410**	.226*	.334**
Updating	.220*	.297**	.204*	.350**	.161	.250**	--	.208*	.051	.217*
Reactivation	.433**	.404**	.510**	.538**	.419**	.538**	.278**	--	.291**	.322**
WRAML	.530**	.337**	.473**	.501**	.395**	.433**	.159	.452**	--	.089
Bridge-IT-global	.247**	.328**	.312**	.504**	.404**	.404**	.259**	.388**	.204*	--

Note. Age = chronological age; TOWRE = Test of Word Reading Efficiency; WJIII-voc = Picture vocabulary subtest from the Woodcock-Johnson III; WJIII-PC = Passage Comprehension subtest from the Woodcock-Johnson III; TORC3-PR = Paragraph Reading subtest from the Test of Reading Comprehension, third edition; WIAT-II-RC = Passage Comprehension subtest from the Wechsler Individual Achievement Test, second edition; Updating = Semantic Updating Task; Reactivation = Semantic Reactivation Task; WRAML = Verbal Working Memory subtest from the Wide Range Assessment of Memory and Learning;

Bridge-IT-global = total accuracy scores (global scores) of the Bridging Inferences Task.

* Correlations are significant at the .05 level (2-tailed test)

** Correlations are significant at the .01 level (2-tailed test)

A total of 12 hierarchical regressions were conducted to test the unique contribution of each predictor of interest (WRAML verbal working memory, semantic updating, semantic reactivation, and Bridge-IT-global accuracy scores) in accounting for variance in each of the three outcome measures of reading comprehension (WJIII-PC, TORC3-PR, and WIAT-II-RC). In all of these regression analyses, age was entered first to control for age related effects, as the experimental measures were not standardized for age. Raw scores from TOWRE (word reading skill) and WJIII-voc (vocabulary knowledge) were entered second to control for word decoding skills and vocabulary knowledge. In the remaining blocks, each of WRAML verbal working memory, semantic updating, semantic reactivation, and Bridge-IT-global accuracy scores were entered last in the regression equation controlling for the other three variables for each of the three measures of comprehension. Raw scores were used for two of the outcome measures (WJIII-PC and TORC3-PR), while weighted raw scores were used for the third outcome measure (WIAT-II-RC). Raw weighted scores were used because each grade read different passages and thus raw scores were not considered comparable.

The hierarchical regressions for the WJIII-PC measure of reading comprehension accounted for 60% of the variance (F (6, 116) = 28.94, $p < .0001$) when working memory, semantic updating, and semantic reactivation were entered in the model and 64.9% of the variance (F (7, 115) = 30.37, $p < .0001$) when Bridge-IT-global accuracy scores were included in the model. The first two steps for each regression were age ($R^2 = .387$, $p < .0001$) and TOWRE word decoding and WJIII-voc vocabulary knowledge ($\Delta R^2 = .157$, $p < .0001$). When WRAML verbal working memory was entered in the third block, before semantic updating and semantic reactivation, it accounted for $\Delta R^2 = .015$ in WJIII-PC scores ($p < .05$; final $\beta = .108$, *ns*). When semantic updating and semantic reactivation were entered after working memory, in the fourth and fifth blocks, they each accounted for $\Delta R^2 = .027$, $p < .01$; final $\beta = .151$, $p < .05$ and $\Delta R^2 = .013$, $p < .05$; final $\beta = .143$, $p < .05$, respectively. When entered in reverse order after working memory, semantic reactivation (Step 4) accounted for $\Delta R^2 = .020$, $p < .05$; final $\beta = .143$, $p < .05$ while semantic updating (Step 5) accounted for $\Delta R^2 = .020$, $p < .05$; final $\beta = .151$, $p < .05$. When semantic updating and semantic reactivation were entered before working memory, in the third and fourth blocks, they accounted for $\Delta R^2 = .029$, $p < .01$; final $\beta = .151$, $p < .05$ and $\Delta R^2 = .019$, $p < .05$; final $\beta = .143$, $p < .05$, respectively. And when working memory was entered last, it did not account for significant variance ($\Delta R^2 = .007$; *ns*, final $\beta = .108$, *ns*). Final beta coefficients revealed that age and WJIII-voc vocabulary knowledge were significant predictors of WJIII-PC scores ($t(121) = 2.37$, $p < .05$, final $\beta = .205$; $t(121) = 4.26$, $p < .0001$; final $\beta = .352$, $p < .0001$, respectively) while TOWRE was not ($t(121) = 1.067$, *ns*; final $\beta = .087$, *ns*). In the fourth regression, when accuracy scores from Bridge-IT global were entered as the final variable (Step 6), after age (step 1; $R^2 = .387$, $p < .001$), TOWRE word decoding and WJIII-voc vocabulary knowledge (Step 2; $\Delta R^2 = .157$, $p < .0001$), working memory (Step 3; $\Delta R^2 = .015$, $p < .05$), semantic updating (Step 4; $\Delta R^2 = .027$, $p < .01$), and semantic reactivation (Step 5; $\Delta R^2 = .013$, $p < .05$) were controlled, it accounted for $\Delta R^2 = .049$ of significant unique variance in WJIII-PC ($p < .0001$; final $\beta = .249$, $p < .0001$). See Appendices 2.A for regression table.

The hierarchical regressions for the TORC3-PR measure of reading comprehension accounted for 30.7% of the variance ($F\,(6, 116) = 8.569$, $p < .0001$) when working memory, semantic updating, and semantic reactivation were entered in the model and 35.1% of the variance ($F\,(7, 115) = 8.882$, $p < .0001$) when Bridge-IT-global accuracy scores were included in the model. The first two steps for the first three regressions were age ($R^2 = .197$, $p < .0001$) and TOWRE word decoding and WJIII-voc vocabulary knowledge ($\Delta R^2 = .067$, $p < .01$). When WRAML verbal working memory was entered in the third block, before semantic updating and semantic reactivation, it accounted for $\Delta R^2 = .022$ in TORC3-PR scores ($p = .056$; final β = .140, *ns*). When semantic updating and semantic reactivation were entered after working memory, in the fourth and fifth blocks, they each accounted for $\Delta R^2 = .000$, *ns*; final β = -.012, *ns* and $\Delta R^2 = .020$, *ns*; final β = .176, *ns*, respectively. When entered in reverse order, semantic reactivation (Step 4) accounted for $\Delta R^2 = .020$, *ns*; final β = .176, *ns* while semantic updating (Step 5) accounted for $\Delta R^2 = .000$, *ns*; final β = -.012, *ns*. When semantic updating and semantic reactivation were entered before working memory, in the third and fourth blocks, they accounted for $\Delta R^2 = .000$, *ns*; final β = -.012, *ns* and $\Delta R^2 = .030$, $p < .05$; final β = .176, *ns*, respectively. And when working memory was entered last, it did not account for significant variance ($\Delta R^2 = .012$; *ns*, final β = .140, *ns*). Examination of final beta coefficients indicated that neither TOWRE or WJIII-voc vocabulary knowledge were significant predictors (t(121) = 1.593, *ns*; final β = .172, *ns*; t(121) = 1.152, *ns*; final β = .125, *ns*; respectively). In the fourth regression, when accuracy scores from Bridge-IT-global were entered as the final variable (Step 6), after age (Step 1; $R^2 = .197$, $p < .0001$), TOWRE word decoding and WJIII-voc vocabulary knowledge (Step 2; $\Delta R^2 = .067$, $p < .01$), working memory (Step 3; $\Delta R^2 = .022$, $p = .056$), semantic updating (Step 4; $\Delta R^2 = .000$, *ns*), and semantic reactivation (Step 5; $\Delta R^2 = .020$, *ns*) were controlled, it accounted for $\Delta R^2 = .044$ of significant unique variance in TORC3-PR ($p < .01$; final β = .235, $p < .01$). See Appendices 2.B for regression table.

The hierarchical regressions for the WIAT-II-RC measure of reading comprehension accounted for 47.3% of the variance ($F\,(6, 116) = 17.34$, $p < .0001$) when working memory, semantic updating, and semantic reactivation were entered in the model and 49.1% of the variance ($F\,(7, 115) = 15.86$, $p < .0001$) when Bridge-IT-global accuracy scores were included in the model. The first two steps for each regression were age ($R^2 = .255$, $p < .0001$) and TOWRE word decoding and WJIII-voc vocabulary knowledge ($\Delta R^2 = .160$, $p < .0001$). When WRAML verbal working memory was entered in the third block, before semantic updating and semantic reactivation, it accounted for $\Delta R^2 = .014$ in WIAT-II-RC scores, *ns*; final β = .084, *ns*. When semantic updating and semantic reactivation were entered after working memory, in the fourth and fifth blocks, they each accounted for $\Delta R^2 = .007$, *ns*, final $\beta = .050$, *ns* and $\Delta R^2 = .037$, $p < .01$; final $\beta = .240$, $p < .01$, respectively. When entered in reverse order, semantic reactivation (Step 4) accounted for $\Delta R^2 = .042$, $p < .01$; final $\beta = .240$, $p < .01$ while semantic updating (Step 5) accounted for $\Delta R^2 = .002$, *ns*; final $\beta = .050$, *ns*. When semantic updating and semantic reactivation were entered before working memory, in the third and fourth blocks, they accounted for $\Delta R^2 = .007$, *ns*, final $\beta = .050$, *ns* and $\Delta R^2 = .046$, $p < .01$; final $\beta = .240$, $p < .01$, respectively. And when working memory was entered last, it did not account for significant variance ($\Delta R^2 = .005$; *ns*; final $\beta = .084$, *ns*). Final beta coefficients revealed that WJIII-voc vocabulary knowledge was a significant predictor of WIAT-II-RC scores ($t(121) = 3.259$, $p < .001$; final $\beta = .309$, $p < .001$) while TOWRE was not ($t(121) = 1.483$, *ns*; final $\beta = .139$, *ns*). In the fourth regression, when accuracy scores from Bridge-IT-global were entered as the final variable (Step 6), after age (Step 1; $R^2 = .255$, $p < .0001$), TOWRE word decoding and WJIII-voc vocabulary knowledge (Step 2; $\Delta R^2 = .160$, $p < .0001$), working memory (Step 3; $\Delta R^2 = .014$, *ns*), semantic updating (Step 4; $\Delta R^2 = .007$, *ns*), and semantic reactivation (Step 5; $\Delta R^2 = .037$, $p < .01$) were controlled, it accounted for $\Delta R^2 = .018$ of significant unique variance in WIAT-II-RC ($p < .05$; final $\beta = .152$, $p < .05$). See Appendices 2.C for regression table.

In summary, working memory did not account for unique variance in any of the outcome measures when updating and reactivation were controlled. The semantic updating task accounted for unique variance in reading comprehension as measured by the WJIII-PC, even after controlling for age, word decoding, vocabulary knowledge, working memory, and semantic reactivation. Semantic reactivation was a unique predictor of performance on the WJIII-PC and WIAT-II-RC, even after controlling for the same variables identified above. Accuracy scores from the experimental Bridging Inference Task (Bridge-IT) accounted for unique variance in performance on all three reading comprehension measures, after controlling for the word-level skills and working memory abilities mentioned above.

Discussion

Demonstrating that somewhat different skills and abilities differentially support performance on different reading comprehension measures is not new; it has been suggested that differences in response format and passage length might account for such findings and also that different reading-related and cognitive measures tap some overlapping, but also some different aspects of the complex process of reading comprehension (Cutting & Scarborough, 2006; Francis et al., 2005; Francis et al., 2006; Keenan et al., 2008; Nation & Snowling, 1997).While these studies focused mainly on the contribution of word decoding skills and listening comprehension to performance on different reading comprehension measures, few studies have explored the relation of specific working memory processes, specific text-level skills such as inference, and word-level decoding and semantic (vocabulary) skills together as predictors of a broad range of standardized measures of reading comprehension. The current study demonstrated that the word-level skills, text-level skills, and cognitive processes differentially predicted reading comprehension achievement on three commonly used standardized reading comprehension outcomes.

In the current study, word decoding was correlated with all three reading comprehension measures but was not significant as a predictor, likely because it was highly correlated with age, as would be expected. Vocabulary knowledge accounted for significant variance across two measures: the WJIII-PC, a measure with relatively short passages, and WIAT-II-RC, a measure with relatively longer passages. Although a measure of verbal working memory span was correlated with all three of the reading comprehension measures it failed to uniquely predict performance on these same measures after the two theoretically-derived working memory components – updating and reactivation - were controlled. In contrast, after controlling for word-level decoding and vocabulary skills, as well as verbal working memory span, semantic updating and semantic reactivation did account for unique variance in performance on two of the three reading comprehension measures. Semantic updating was a significant unique predictor of performance on the WJIII-PC and semantic reactivation was a significant, unique predictor of performance on both the WJIII-PC and the WIAT-II-RC. In addition, a text-level skill, bridging inference making, uniquely predicted performance on all three reading comprehension measures, even when decoding, vocabulary, and working memory were controlled

Working Memory and Working Memory Components: Semantic Updating and Reactivation

Working memory component measures (i.e., updating and reactivation) were significantly correlated with the three reading comprehension measures (with the exception of semantic updating and the TORC3-PR) but whether these components uniquely predicted reading comprehension with word decoding, vocabulary knowledge, and verbal working memory span in the models, was dependent on the reading comprehension measure in question.

As predicted, semantic updating was a unique predictor of performance on the short-passage (one or two sentences) cloze task (WJIII-PC) but did not uniquely predict performance on either of the longer-passage multiple-choice or open-ended format tasks (TORC3-PR and WIAT-II-RC). This finding makes sense given the suggestion made earlier that the WJIII-PC taps into the mental model *building* aspect of reading comprehension, a process that would continuously involve updating or revising the content of working memory to include newly presented information such as general knowledge (i.e., word meanings). It may be that semantic updating did not predict performance on the TORC3-PR and WIAT-II-RC because, as discussed above, both measures may be measuring the readers' ability to build a mental model and then *reflect* upon that model in order to respond to questions.

Semantic reactivation predicted unique variance on the WJIII-PC and the WIAT-II-RC but not the

TORC3-PR. Because the cloze task on the WJIII-PC often required the reader to provide a missing word at the end of the sentence or short passage, the reader may have reactivated early portions of the sentence or text in order to use context to aid in the completion of the item. Consistent with predictions, semantic reactivation predicted children's performance when answering open-ended questions about longer passages (10 sentences or more) they read independently (WIAT-II-RC), even when working memory was controlled. This finding suggests that children reactivated or brought the specific information needed to correctly respond to questions from the passage back into working memory when cued to do so by the examiner's question (Albrecht & O'Brien, 1993; Gernsbacher, 1990; Kintsch, 1991; Linderholm et al., 2004). Although the text was made available to children, they may have chosen to not refer back to the text and instead proceeded to reactivate and reflect upon the mental model held in working memory, placing a heavier burden on the reactivation component in working memory

Working memory comprises various components (e.g., updating, Palladino et al., 2001; reactivation, Linderholm et al., 2004; suppression, Gernsbacher & Faust, 1991, Pimperton & Nation, 2010; attentional focus, Engle, 2002; Unsworth & Engle, 2007;) that are strategically engaged throughout the comprehension process. The finding that the measure of working memory span did not predict unique variance on any reading comprehension measure after updating and reactivation were controlled, suggests that the processes in working memory that account for the relation of verbal working memory span with some of the reading comprehension measures is synonymous with the processes captured by the updating and/or reactivation tasks. The current study's findings suggest that it is updating and reactivation processes that are linked to reading comprehension.

Bridging Inferences

A text-level skill important for maintaining narrative coherence, bridging inference making, accounted for significant unique variance on all three comprehension measures over and above word-level decoding and semantic skills, verbal working memory span, and specific working memory processes. This finding is consistent with previous findings in the literature (Cain & Oakhill, 1999; Cain et al., 2004; Johnston, 2006; Pike et al., 2010) and suggests that the ability to integrate explicitly stated ideas within text might be necessary for deriving meaning. Bridging inferences can help readers decipher word meanings from context when the reader does not know the meaning of an unfamiliar word (Cain et al., 2004). Although cognitive processes, such as working memory, updating, and reactivation, support inferencing (Chapter 2), they did not fully account for the relation with comprehension given that bridging inferences accounted for unique additional variance in reading comprehension. Other factors that may influence one's ability to make bridging inferences may include knowledge about when to make a bridging inference (e.g., reading strategies; Cain & Oakhill, 1999).

Conclusions, Limitations, and Future Directions

This is one of the first studies to demonstrate relations of specific working memory processes, updating and reactivation, and bridging inferences while controlling for word decoding and vocabulary knowledge to three different comprehension measures together in one study. However, the results and implications of this study should be interpreted with caution. This study included a relatively small sample size (123 children) across six grades. It would be ideal to replicate the current findings with a larger sample of children. As discussed above, readers are often required to go beyond the text itself and make connections between general knowledge about the world and the ideas presented in text (Barnes, Dennis, & Haefele-Kalvaitis, 1996; Cain, Oakhill, Barnes, & Bryant, 2001). Readers often construct a situational model that exists in time and space, involves cause and effect relationships, and has characters that engage in goal directed behaviour (Kintsch & Kintsch, 2005; Zwann & Radvansky, 1998). It would be interesting to explore whether other cognitive processes such as visual working memory would predict performance on tasks that may encourage readers to form situational models, such as the TORC3-PR or WIAT-II-RC.

In conclusion, the contributions of the current study to understanding the factors that influence performance on several commonly used reading comprehension tests are that word-level decoding and semantic skills, text- or discourse-level skills such as bridging inference, and general purpose working memory processes, such as updating and reactivation, support reading comprehension achievement, but the contribution of these skills varies somewhat as a function of the specific achievement measure used to assess reading comprehension.

References

Albrecht, J. E., & O'Brien, E. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(5), 1061 – 1070.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Clarendon Press.

Barnes, M. A., Dennis, M., & Haefele-Kalvaitis, J. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inferencing in children from six to fifteen years of age. *Journal of Experimental Child Psychology, 61*(3), 216-241.

Barnes, M. A., Faulkner, H., Wilkinson, M., & Dennis, M. (2004). Meaning construction and integration in children with hydrocephalus. *Brain and Language, 89*, 47 - 56.

Barnes, M. A., Johnston, A. M., & Dennis, M. (2008). Comprehension in a neurodevelopmental disorder, spina bifida Myelomeningocele. In K. Cain, K., & J. Oakville. (Eds). *Children's comprehension problems in oral written language: A cognitive perspective. Challenges in language and literacy.* (pp. 193 – 217). New York, NY, US: Guilford Press.

Bowyer-Crane, C., & Snowling, M. J. (2005). Assessing children's inference generation: What do tests of reading comprehension measure? *British Journal of Educational Psychology, 75*, 189 – 201.

Brown, V., Hammill, D. D., & Wiederholt, J. L. (1995).*The Test of Reading Comprehension-Third Edition*. Austin, TX: PRO-ED.

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing, 11*(5-6), 489-503.

Cain, K., Oakhill, J. V., Barnes, M. A., Bryant, P. E. (2001). Comprehension skill, inference-making ability, and the relation to knowledge. *Memory & Cognition, 29*(6), 850-859.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31-42.

Cain, K., Oakhill, J., &Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language*, *30*, 681– 694.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96*, 671–681.

Carretti, B., Borella, E., Cornoldi, C., & DeBeni, R. (2009). The role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences, 19*, 246 – 251.

Carretti, B., Cornoldi, C., De Beni, R., Romano, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology, 91*, 45-66.

Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J.,Willcutt, E., & Olson, R. K. (2012). Predicting Word Reading and Comprehension With Executive Function and Speed Measures Across Development: A Latent Variable Analysis. *Journal of Experimental Psychology: General, 141*(3), 470 – 88.

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10*, 277 - 299.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19-23.

Francis, D. J., Fletcher, J. M., Catts, H. W., & Tomblin. (2005). In *Children's reading comprehension and assessment*. Eds. Paris, Scott G., Stahl, Steven A. (369 – 394. Mahwah, NJ, US: Lawrence Erlbaum.

Francis, D. J., Snow, C. E., August, D., Carlson, C. D., Miller, J., & Iglesias, A. (2006). Measures of reading comprehension: A latent variable analysis of the diagnostic assessment of reading comprehension. *Scientific Studies of Reading, 10*(3), 301 – 322.

Field, A. (2005). *Discovering statistics using SPSS.*(2nd ed.). London, UK: Sage.

Gernsbacher, M. A. (1990).*Language comprehension as structure building.* Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Gernsbacher, M.A. & Faust, M.E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition*, *17*, 245–262.

Gernsbacher, M. A., Robertson, R. R., Palladino, P., & Werner, N. K. (2004). Managing mental representations during narrative comprehension. *Discourse Processes, 37*(2), 145 – 164.

Gough, P. & Tunmer, W. (1986).Decoding, reading, and reading disability. Remedial and Special Education, 7, 6–10.

Hulme, C., & Snowling, M. J. (2011). Children's reading comprehension difficulties: Nature, causes, and treatments. *Current Directions in Psychological Science, 20*(3), 139 - 142.

Johnston, A. M. (2006). *The role of attentional control in reading comprehension*. Unpublished master's thesis, University of Guelph, Guelph, Ontario, Canada.

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*(3), 281 - 300.

Kintsch, W. A. (Ed.). (2002). *The role of knowledge in discourse comprehension: A construction-*

integration model. Cambridge, MA, US: MIT Press.

Kintsch, W., &Kintsch, E. (2005). Comprehension. In S. G. Paris and S. A. Stahl (Eds.), *Children's Reading Comprehension and Assessment* (pp. 71 - 92). New York: Lawrence Erlbaum.

Kirby, J. R., & Savage, R. S. (2008). Can the simple view deal with the complexities of reading? *Literacy, 42*(2), 75 – 82.

Linderholm, T., Gernsbacher, M. A., Van den Broek, P., Neninde, L., Robertson, R. R. W., & Sundermier, B. (2004). Suppression of story character goals during reading. *Discourse Processes, 37*(1), 67 – 78.

Linderholm, T., Virtue, S., Tzeng, Y., and van den Broek, P. (2004). Fluctuations in the availability of information during reading: Capturing cognitive processes using the landscape model. *Discourse Processes, 37*(2), 165 – 186.

Nation, K., & Snowling, M. (1997). Assessing reading difficulties: the validity and utility of current measures of reading skill. *British Journal of Educational Psychology, 67*, 359–370.

Nation, K., & Snowling, M. J. (2000). Factors influencing syntactic awareness skills in normal readers and poor comprehenders. Applied Psycholinguistics, 21, 229 - 241.

Oakhill, J. (1993). Children's difficulties in reading comprehension. *Educational Psychology Review. Special Issue: European Educational Psychology, 5*(3), 223-237.

O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., &Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1200 – 1210.

Palladino, P., Cornoldi, C., DeBeni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory & Cognition, 29*(2), 344-354.

Perfetti, C., Yang, C-L., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology, 22,* 303 – 318.

Pike, M. M., Barnes, M. A., & Barron, R. W. (2010). The role of illustrations in children's inferential comprehension. *Journal of Experimental Child Psychology, 105*, 243 – 255.

Pimperton, H., & Nation, K. (2010). Suppressing irrelevant information from working memory: Evidence for domain-specific deficits in poor comprehenders. *Journal of Memory and Language, 62,* 380 – 391.

Schmalhofer, F., McDaniel, M. A., & Keefe, D. (2002). A unified model for predictive and bridging inferences. *Discourse Processes, 33*(2), 105-132.

Sheslow, D. V., & Adams, W. V. (2003). Wide range assessment of memory and learning-second edition. Wilmington, DE: Wide Range.

Snow, C. (2002). RAND Reading Study Group. Reading for understanding. Santa Monica, CA: RAND.

Torgensen, J. K., Wagner, R. K. & Rashotte, C. A. (1999).*Test of Word Reading Efficiency* (*TOWRE*). Austin, TX: Pro-Ed.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*(1), 104 – 132.

Wolf, M., & Katzir-Cohen, T. (2001).Reading fluency and its intervention. *Scientific Studies of Reading, 5*, 211 – 238.

van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes, 39*(2 & 3), 299 - 316.

van den Broek, P., Young, M., Tzeng, Y., &Linderholm, T. (Eds.). (1999). *The landscape model of reading: Inferences and the online construction of memory representation.* Mahwah, NJ, US: Lawrence Erlbaum Associates, Publishers.

Wechsler, D. (2002). *Manual for the Wechsler Individual Achievement Scale, 2nd Edition.* Toronto: The Psychological Corporation.

Woodcock, R., Mather, N., & McGrew (2001).*Woodcock-Johnson III – Test of Achievement.* Chicago: Riverside.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162-185.

Appendix 2.A

Table 2.A.1

Summary of Hierarchical Regression Predicting Reading Comprehension Achievement on the WJIII-PC (N = 123)

Outcome	Step	Variable	ΔR^2	B	SE B	β
WJIII-PC						
	1	Age	.387***	.417	.176	.205*
	2	Word Decoding	.157***	.028	.026	.087
		Vocabulary		.361	.085	.352***
	3	Working memory	.015*	.069	.047	.108
	4	Updating	.027**	.157	.065	.151*
	5	Reactivation	.013*	.070	.036	.143*
	4	Reactivation	.020*	.070	.036	.143*
	5	Updating	.020*	.157	.065	.151*
	3	Updating	.029**	.157	.065	.151*
	4	Reactivation	.019*	.070	.036	.143*
	5	Working memory	.007	.069	.047	.108
	3	Working memory	.015*	.070	.044	.110
	4	Updating	.027**	.123	.062	.118*

5	Reactivation	.013*	.038	.034	.077
6	Bridging inferences	.049***	.299	.074	.249***

*p** = 0.05; *p*** = 0.01; *p**** = 0.001

Appendix 2.B

Table 2.B.1

Summary of Hierarchical Regression Predicting Reading Comprehension Achievement on the TORC3-PR (N = 123)

Outcome	Step	Variable	ΔR^2	B	SE B	β
TORC3-PR						
	1	Age	.197***	.433	.434	.114
	2	Word Decoding	.067**	.102	.064	.172
		Vocabulary		.241	.209	.125
	3	Working memory	.022	.167	.116	.140
	4	Updating	.000	-.024	.161	-.012
	5	Reactivation	.020	.161	.088	.176
	4	Reactivation	.020	.161	.088	.176
	5	Updating	.000	-.024	.161	-.012
	3	Updating	.000	-.024	.161	-.012
	4	Reactivation	.030*	.161	.088	.176
	5	Working memory	.012	.167	.116	.140
	3	Working memory	.022	.169	.112	.142
	4	Updating	.000	-.084	.158	-.043

5	Reactivation	.020	.104	.088	.114
6	Bridging inferences	.044**	.529	.190	.235**

*p** = 0.05; *p*** = 0.01; *p**** = 0.001

Appendix 2.C

Table 2.C.1

Summary of Hierarchical Regression Predicting Reading Comprehension Achievement on the WIAT-II-RC (*N* = 123)

Outcome	Step	Variable	ΔR^2	B	SE B	β
WIAT-II-RC						
	1	Age	.255***	1.047	1.409	.074
	2	Word Decoding	.160***	.307	.207	.139
		Vocabulary		2.211	.679	.309**
	3	Working memory	.014	.375	.375	.084
	4	Updating	.007	.360	.523	.050
	5	Reactivation	.037**	.817	.285	.240**
	4	Reactivation	.042**	.817	.285	.240**
	5	Updating	.002	.360	.523	.050
	3	Updating	.007	.360	.523	.050
	4	Reactivation	.046**	.817	.285	.240**
	5	Working memory	.005	.375	.375	.084
	3	Working memory	.014	.381	.370	.086
	4	Updating	.007	.216	.520	.030

5	Reactivation	.037**	.680	.289	.200*
6	Bridging inferences	.018*	1.273	.625	.152*

*p** = 0.05; *p*** = 0.01; *p**** = 0.001

Chapter 4: General Discussion

This dissertation includes two studies that explored related but distinct questions about reading comprehension. The first study, *Bridging Inferences and their Relation to Working Memory Processes in Children in Grades Three through Eight* was an in depth investigation of grade-related effects on bridging inferences as a function of text distance (integrating information between contiguous sentences or separated by larger chunks of text) and domain knowledge (information based on cause and effect relationships or based on character goals, motivations, and desires). This is the first study to investigate the effects of text distance and knowledge domain on bridging inferences in the same study with any population. The second study *The Relation of Working Memory Processes and Bridging Inferences to Reading Comprehension in Children in Grades Three through Eight* investigated whether word decoding, vocabulary, and text-level bridging inferences accounted for significant variance on three commonly used reading comprehension measures. In addition, Study 1 investigated specific working memory components (semantic updating and semantic reactivation) in relation to the text-level skill of bridging inferences and Study 2 investigated these same working memory processes in relation to general reading comprehension achievement.

Bridging Inferences.

Study one demonstrated that school aged children become more accurate when making bridging inferences across grades. Both increasing textual distance between to-be-integrated information and the use of affective themes made inference-making more difficult compared to when information was closer together in the text, or when inferences were made using concrete-causal information. These findings extend our knowledge of bridging inferences in children given that this is the first study to explore the effects of text distance and domain knowledge in the same study. Performance on bridging inference predicted performance on all three standardized reading comprehension tasks, providing further support for the unique contribution of bridging inferences in reading comprehension.

Working Memory, Semantic Updating, and Semantic Reactivation.

Bridging inferences. Study one is the first study to look at the cognitive resources needed to make bridging inferences as a function of text distance and domain knowledge. Semantic reactivation (the process of bringing previously processed information back into working memory) predicted variance when the information needing integration was farther apart in the text, regardless of knowledge domain. While neither working memory component accounted for unique variance in the near inferencing conditions, reactivation and updating (the process of efficiently updating the contents of working memory) predicted performance on the far-concrete inferencing condition. These findings suggest that not only are updating and reactivation important for bridging inferences in children, but the combination of cognitive processes involved in making bridging inferences differ depending on text distance between to-be-integrated information and different types of knowledge.

Reading comprehension. The second study demonstrated that the working memory processes involved in comprehension differ among standardized comprehension measures. Specifically, semantic updating and semantic reactivation predicted unique variance in performance on the WJIII-PC, a cloze task with relatively short passages that was argued to measure the ability *to build* a mental model. Semantic reactivation predicted unique variance in performance on the WIAT-II-RC, a measure with an open-ended response format and relatively longer passages that were purported to measure the ability to build and then *reflect* upon the mental model. These findings suggest that the working memory processes that accounted for relations with reading comprehension may include updating and reactivation.

In conclusion, grade, text, and domain knowledge effects were found to affect children’s ability to make bridging inferences. Bridging inferences uniquely contributed to reading comprehension on a variety of reading comprehension tasks, highlighting its importance in reading comprehension achievement in school-aged children. Different working memory processes, semantic updating and semantic reactivation, were uniquely important for a variety of comprehension tasks that varied in length and response format, suggesting that readers engage different cognitive processes depending on the demands for constructing meaning while comprehending text.

Limitations and Future Directions

Despite the findings outlined above, the results and implications discussed in this dissertation should be interpreted with caution. First, both studies should be replicated with larger sample sizes. Second, continuing to explore which cognitive processes support performance on different comprehension measures may provide those who use different comprehension tools (e.g., for cognitive assessment and intervention development) with information that would allow them to strategically choose tools to answer particular questions. Additionally, having a better understanding about what different tools actually measure may provide alternative explanations for findings with said tools. Third, the Bridging Inferences Task has been used in previous studies and has shown similar text distance findings with different samples; however, the task could be further developed to explore distance effects across larger chunks of text while also further investigating inference-making with different types of domain knowledge. Extending the text may allow findings to be more generalizable to situations that children and adolescents more commonly encounter while reading. Fourth, the Semantic Updating Task and Semantic Reactivation Task are experimental tasks that may require further development to increase reliability, validity, and sensitivity. In Chapter 2, it was not possible to tell whether suppression was uniquely important for bridging inferences. Further research could investigate the role of suppression and reactivation as well as other cognitive processes such as attentional shifting in bridging inferences in the same study.

References

Barnes, M. A., Dennis, M., & Haefele-Kalvaitis, J. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inferencing in children from six to fifteen years of age. *Journal of Experimental Child Psychology, 61*(3), 216-241.

Barnes, M. A., Faulkner, H. Wilkinson, M., & Dennis, M. (2004). Meaning construction and integration in children with hydrocephalus. *Brain and Language, 89*, 47 - 56.

Barnes, M. A., Johnston, A. M., & Dennis, M. (2008). Comprehension in a neurodevelopmental disorder, spina bifida Myelomeningocele. In K. Cain, K., & J. Oakville. (Eds). *Children's comprehension problems in oral written language: A cognitive perspective. Challenges in language and literacy.*(pp. 193 – 217). New York, NY, US: Guilford Press.

Castell, M. A., & Simpson, G. B. (1991). Textual coherence and the development of inferential generation skills. *Journal of Research in Reading, 14*(2), 116-129.

Cain, K., Oakhill, J. V., &Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language, 30*, 681 - 694.

Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J., Willcutt, E., & Olson, R. K. (2012, February 20). Predicting Word Reading and Comprehension With Executive Function and Speed Measures Across Development: A Latent Variable Analysis. Journal of Experimental Psychology: General. Advance online publication. doi: 10.1037/a0027375.

Gernsbacher, A. M. (1997). Attentuating interference during comprehension: The role of suppression. *The Psychology of Learning and Motivation, 37*, 85 - 104.

Johnston, A. M., Barnes, M. A., & Desrochers, A. (2008, June). Bridging Inferences in School Aged

Children. Poster presented at the 39th Annual Convention of the Canadian Psychology Association (CPA), Halifax, NS.

Oakhill, J., Cain, K., & Bryant, P. W. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18*(4), 443 – 468.

Pike, M. M., Barnes, M. A., & Barron, R. W. (2010). The role of illustrations in children's inferential comprehension. *Journal of Experimental child Psychology, 105*, 342 – 255.

Pimperton, H., & Nation, K. (2010). Suppressing irrelevant information from working memory: Evidence for domain-specific deficits in poor comprehenders. *Journal of Memory and Language, 62*, 380 – 391.

Gernsbacher, M. A., & Shlesinger, M. (1997). The proposed role of suppression in simultaneous interpretation. *Interpreting, 2* (1/2), 119 - 140.

Shears, C., Miller, V., Ball, M., Hawkins, A., Griggs, J., & Varner, A. (2007). Cognitive demand differences in causal inferences: Characters' plans are more difficult to comprehend than physical causation. *Discourse Processes, 43*(3), 255 – 278.

Appendix A

Semantic Updating and Reactivation Tasks

The Semantic Updating Task. To test if there were differences in semantic updating between grades, one-way ANOVA was conducted on updating scores. A main effects of grade was not found (F (5, 122) = 2.12; *ns*).

Table A.1

Means (standard deviations) for the Semantic Updating Task

Grade 3 n = 19	Grade 4 n = 25	Grade 5 n = 23	Grade 6 n= 21	Grade 7 n = 21	Grade 8 n = 19
6.63 (2.98)	7.08 (3.60)	7.30 (3.18)	7.76 (2.88)	8.86 (3.85)	9.42 (3.34)

The Reactivation Task. To test if there were differences in semantic reactivation between grades, a one-way ANOVA was conducted on reactivation scores. A main effect of grade was found for semantic reactivation (F (5, 122) = 7.50; $p < 0.0001$). Tukey post-hoc tests ($p< .05$) revealed that the grade three total reactivation scores were significantly lower than the total reactivation scores of grades six through eight and grade four and five scores were significantly lower than grade eight scores. See Table A.2 for means (standard errors).

Table A.2

Proportional Means (standard error) for the Semantic Reactivation Task by Grade

Grade 3 n = 19	Grade 4 n = 25	Grade 5 n = 23	Grade 6 n = 21	Grade 7 n = 21	Grade 8 n = 19
4.93 (1.13)	5.71 (1.17)	5.87 (1.26)	6.27 (1.15)	6.69 (1.60)	7.08 (.89)