Influence of Working Memory Capacity and Reading Purpose on Young Readers’ Text Comprehension

Zeynep Ozlem Cankaya, B.Sc.
Department of Educational and Counselling Psychology
McGill University, Montreal
August, 2007

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Arts in Educational Psychology, Major in Applied Cognitive Science.

© Zeynep Ozlem Cankaya 2007
Abstract

Reading comprehension processes are assumed to be influenced by reading purpose and working memory capacity (WMC). However, it is still unknown how these factors affect comprehension processes in young readers. The aim of this study was to explore whether cognitive processes varied as a function of reading purpose (test versus game) and WMC (high versus low) in young readers. The 39 participants completed the Working Memory Test Battery for Children (WMTB-C), a verbal protocol and a free-recall task. Separate ANOVAs on cognitive processes response categories detected medium effect sizes. In the free recall task, test condition readers exhibited more paraphrasing and recalled more idea units than readers in the game condition. In the verbal protocol task, readers in the game condition uttered more evaluative comments than in the test condition. Furthermore, low WMC readers produced more predictive inferences than the high WMC group. Possible contributions of reading purpose and WMC to text comprehension for educational practice were discussed.

Résumé de Recherche

Les processus cognitifs impliqués dans la compréhension de textes sont influencés par le but de la lecture et la capacité de mémoire de travail (CMT). Toutefois, nous ignorons toujours comment ces facteurs influencent la lecture chez les jeunes lecteurs. Le but de cette étude était de vérifier si les processus cognitifs varient en fonction du but de la lecture (test versus jeu) et de la capacité de la mémoire de travail (faible versus élevée) chez les jeunes enfants. Les trente-neuf participants de l’étude ont complété le Working Memory Test Battery for Children (WMTB-C), un protocole verbal et une tâche de rappel libre. Les analyses statistiques comparant les différentes catégories de processus
cognitifs ont révélé des effets de taille moyenne. Pour le rappel libre, les lecteurs ont paraphrasé davantage et ont mémorisé plus de groupes d’idées dans la condition test que la condition jeu. Lors du protocole verbal, les lecteurs de la condition jeu ont fait plus de commentaires évaluatifs que dans la condition test. Finalement, les enfants ayant une CMT plus faible ont prononcé plus d’inférences de prédiction que ceux ayant une CMT plus élevée. La contribution des processus cognitifs et de la CMT à la compréhension de lecture dans un contexte éducatif fut considérée.

Acknowledgements

First and foremost, I would like to thank my parents and my brother, who valued my dreams and encouraged me towards realizing them. To my best friend Wanda Vreden for her love and understanding as I worked towards the completion of my master’s thesis, who has been a wonderful shoulder, whenever I needed even at the times she was physically away.

The completion of my thesis would not be possible without the support, and constructive advice of my supervisor, Dr. Robert Savage. I would like to express appreciation to him, who has scaffolded my learning process by sharing his expertise and endless enthusiasm. I also would like to use this opportunity to thank my inspirational mentor Dr. Jeff Bisanz, who changed my direction in life profoundly.

Whatever I say, it will be insufficient to express thanks to various people: To our research team at McGill University for all of their hard work and support on this project, particularly to Louise Deault for her astute advice on methodology; Marc Garellek, and Sherry Smith for being generous and patient with their help in collecting the data collection and scoring them. Special thanks to Lisa-Marie Sauvé for translating the
abstract into French and to dear colleague and friend Krista Redden, who unconditionally offered me her expertise and warmth. My final debt of gratitude goes to good friends Ozgur Oner, Ozgur Parlak, Ozlem Erten, Gunter Soydanbay, Gulce Kucur, Tuba Eren, Meltem Sariahmetoglu for their sustained support and encouragement in the course of my graduate degree.

Finally, I would like to thank all of the staff and students at the Riverside School Board who assisted me in the completion of this project. Particularly, I am grateful to the principle of St. Mary’s Elementary School, Ms. Julia Pavlopooulos, and the school staff for their assistance and kindness.
# Table of Contents

Abstract.................................................................................................................................................2
Résumé.....................................................................................................................................................2
Acknowledgements.................................................................................................................................3
Table of Contents...................................................................................................................................5
List of Tables...........................................................................................................................................5
Introduction .............................................................................................................................................7
Review of Literature ..............................................................................................................................9
Aims of the Current Study ....................................................................................................................23
Method ....................................................................................................................................................24
  Participants .........................................................................................................................................24
  Materials ............................................................................................................................................25
  Measures .............................................................................................................................................26
  Procedure ...........................................................................................................................................33
  Data Scoring .......................................................................................................................................37
Results ....................................................................................................................................................40
Discussion .............................................................................................................................................50
  Limitations .........................................................................................................................................57
References ..............................................................................................................................................62
Appendices ..........................................................................................................................................73
  Appendix A. Text: The ancient Game of Go ......................................................................................73
  Appendix B. The Visual Comprehension Questions of Go Game ....................................................76
  Appendix C: The Verbal Comprehension Questions of Go Game ..................................................79
Appendix D. Ethics Approval Form……………………………………………………80

List of Tables

Table 1: Sample Characteristics: Standard Scores of Standardized Measurements……68
Table 2: ANOVA Results for the Total Sample for Verbal Protocol Task………………68
Table 3: ANOVA Results for the Total Sample for Free Recall Task …………………69
Table 4: ANOVA Results for the Total Sample for Free Recalled idea units …………..71
Table 5: ANOVA Results for the Total Sample for Comprehension-Application Tasks 71
Introduction

Reading comprehension is rather a hard task among all other components of reading coherence. Even fluent readers sometimes have problems with understanding what they are reading about during the course of reading. What makes text comprehension difficult is a complicated question to answer. The task analysis of reading comprehension is not straightforward because numerous components contribute individually while the person reads (Dunlosky et al., 2007). Complex cognitive tasks require people to maintain a large amount of incoming information. Furthermore, reading is one of those complex tasks which simultaneously puts a lot of demand on both storage and processing function of memory. Thus, there is a growing recognition of the importance of working memory in the processes of reading. Evidence suggests that there is a clear link between working memory and text comprehension (Daneman & Carpenter, 1980; Baddeley, 1990).

When we read, we do it for different purposes: Sometimes for preparing for a test, or sometimes just for playing a game. In a study Lorch, Lorch, and Klusewitz (1993) asked college readers to distinguish different reading situations. The results suggested that the readers identified at least 10 distinct categories, such as personal choice reading and school reading. The participants also suggested that the cognitive demand is different for each reading situation.

According to the current models of reading, the reading comprehension is affected by working memory capacity and reading purpose in readers (Linderholm & van den Broek, 2002). A few studies explored the role of working memory and reading comprehension in adults, but none in young readers. How does reading comprehension differ when young readers are given different reading purposes? In particular, the aim of
the present study is to evaluate the impact of reading purpose on reading comprehension within the context of high and low working memory capacities. This paper will review current research on the experiential and cognitive factors involved in the complex relationship between working memory and the cognitive processes to comprehend text in young readers, and discuss the potential effects of providing reading purpose in order to elicit required or necessary comprehension processes.

Previous research explored the effect of WM and reading purpose on reading comprehension in adult readers, where the reading processes is automatized. However, no work has explored the combination of these two factors in young readers’ reading comprehension. Savage, Lavers, & Pillay (2007) point out that further work is needed to specify the nature of WM problems and reading comprehension. An important issue that has not been directly touched upon, whether advances in understanding the link between reading purpose and reading comprehension can be put into a practical use. The results of this study might be essential in understanding how particular reading goals could be provided in order to elicit desired cognitive processes.

For that reason, in this present study, the aim is to explore the contribution of WMC and reading purpose to young reader’s text comprehension in an applied context.

The objectives of the study are

(1) to investigate the effect of reading purpose and WM on text compression,

(2) to explore the depth of reading comprehension processes under different reading purposes,

(3) and to investigate these in an applied context to obtain ecological validity.
The first section of this paper will briefly summarize the models of working memory with particular attention to Baddeley and Hitch’s account of working memory. In this section, the available methods for assessing working memory capacity will be discussed. The next section will review the relationship between working memory and reading comprehension with a particular focus on how reading comprehension differs in regards to individual differences in working memory capacity. The third section of the paper will review the literature on the effects of reading purpose on cognitive processes. The link between different cognitive processes involved in reading, when the readers are provided with different reading purposes, will be discussed. Finally, the paper will explore the effects of different contexts on reading comprehension.

**Literature Review**

*Baddeley’s Working Memory Model*

Working memory is assumed to be a dynamic system with a capacity to store information while engaging in other cognitively demanding activities (Savage, Lavers, & Pillay, 2007). Baddeley and Hitch (1974) proposed a three component model of Working Memory (WM), which has been one of the most influential accounts for how we process the incoming information. In the original model of Working Memory, there are three main subsystems: (1) The phonological loop is for maintaining and rehearsing speech based, verbal and acoustic information (Baddeley, 2000); (2) the visual–spatial sketch pad is for storing and manipulating incoming visual, spatial and kinesthetic component information; and (3) the central executive for the control and regulation of attention system and cognitive processes. This system has power over the activation of long-term memory (Baddeley, 1998), as well as shifting between tasks or retrieval strategies.
(Baddeley, 1996), and it harmonizes multiple tasks (e.g., Baddeley & Della Sala, 1996). Baddeley (2000) has revised our understanding of the set up of working memory by proposing another component, the episodic buffer, which is a multi-dimensional code from the two subsystems and long-term memory. This component provides a temporary crossing point between the two slave systems and the central executive, which is responsible for binding information into coherent episodes (Baddeley, 2000). This system as a whole provides the resources needed to retrieve and maintain information during cognitive processing (Baddeley, 1990).

This model is critical in understanding the reading comprehension as language comprehension depends on verbal part of working memory, in particular phonological loop is crucial. Furthermore, the central executive contributes to comprehension in a general, but crucial sense (Gathercole & Baddeley, 1993). Central executive has been associated closely with complex cognitive tasks (Boduroglu et al., 2007). Understanding individual differences of working memory would benefit how we generate new instruction methods for improving reading comprehension.

Another influential theoretical approach to explaining working memory concept is suggested by Ericsson and Kintsch (1995) and called Long-Term Working Memory. As opposed to Baddeley and Hitch’s WM Model, Long-Term Working Memory is a construct within the Long Term Memory structure. According to their assumptions we use memory in most everyday tasks. According to Ericsson and Kintsch (1995), the information is stored in the long term memory during reading. For instance, reading necessitates the reader to maintain the registered information in memory with several
chunks. We accumulate most of what we read in long-term memory, linking them all together through retrieval clues. Since, the limited chunk capacity of our memory system would be exceeded after a few sentences. This would result in not being able to understand the multifaceted relations between thoughts expressed in a novel or a scientific text. We need to hold only a few concepts in working memory, which serve as cues to retrieve everything associated to them by the retrieval structures (Ericsson & Kintsch, 1995). The skilled readers’ representations are more accessible than the less skilled readers because of the former group’s effective retrieval structures. Kintsch (1998) describes this idea with a metaphor: Good readers do not have a larger box to fit things in for temporary storage; they are only more skilled in putting things into long term storage and retrieving them. In Ericsson and Kintsch’s account, LT-WM serves to update current stored information by accessing relevant existing information, thus operating like a retrieval cue (Savage, Lavers, & Pillay, 2007). However, while influential in current work in the adult literature, Baddeley’s multi-component WM modal has been explored more in the children’s literature.

The pioneering study of Daneman and Carpenter (1980) explored individual differences in the capacity of working memory requiring simultaneous storage and processing of information. In their study, they created a task which taps both storage and processing functions of WM proposed by Baddeley and Hitch. They named this task the reading span test, in which the participant is presented a series of sentences. For instance,

“When at last his eyes opened, there was no gleam of triumph, no shade of anger.”

“The taxi turned up to Michigan Avenue, where they had a clear view of the lake.”
Next, the participant judges the semantic properties of sentences while remembering the last word of each sentence in sequence (In the example above, the correct response is: “anger” and “lake”), after the introduction of the last sentence in the list of few (Daneman & Carpenter, 1980). The importance of this study is that the reading span test was the best measure of the working memory available or functional during reading which would assess both the storage and processing aspects of this working memory system.

The span defines the number of sentences, accurately processed and remembered (Baddeley, 2007). They found out that the readers who were less skilled allocated more capacity or resources to processing the sentences, resulting in less capacity to offer for maintaining the incoming data. On the other hand, skilled readers had more capacity to devote to data maintenance because their processing was more efficient. This measure highly correlated with reading comprehension scores of the participants.

Although at the time Daneman and Carpenter’s working memory span task was a revolutionary tool, it has not gone without serious criticisms: Waters and Kaplan (1996) for example claimed that this task originally intended to tax both processing and storage functions of working memory; however, it only measures the storage function. Furthermore, in their extensive review of what we know and do not know about working memory, Savage et al. (2007) claimed that because the complex span task requires explicit and effortful processing demands, it may not be most appropriate to measure comprehension in naturalistic settings.

According to Pickering and Gathercole (2001), in order to provide a reliable assessment of working memory abilities, different measures have to be utilized. The intersection of different measures reduces the chance of concluding erroneously. In a large-scale study on the structure of working memory in children, Gathercole et al.
(2004) found that the three-component model of working memory best fit the data and was supported in 6- and 7-year-olds, 8- and 9-year-olds, 10- to 12-year-olds, and 13- to 15-year-olds.

The pattern of findings, using different working memory measures have been hard to interpret because tasks and procedures vary vastly in both adult and children’s reading comprehension literature (Savage, Lavers, & Pillay, 2007).

In the existing WM literature there is a varying set of different measures, such as complex span task, simple span task etc. This makes it difficult to compare reliability across studies that used different span tasks, scoring procedures and samples (Seigneuric & Ehrlich, 2005). Therefore, the text comprehension measures should be more comprehensive in order to understand how readers adapt their reading processes. Pickering and Gathercole (2001) devised a set of complementary measures called Working Memory Test Battery for Children (WMTB-C). This comprehensive battery has benefited the literature, as it intends to measure the three components of Working Memory modal. Therefore, this robust, wide-ranging assessment tool of three components can lead a unified understanding in the role of working memory by producing greater consistency in the literature than has been right now (Savage, Lavers, & Pillay, 2007).

Working Memory and Reading Comprehension

Reading comprehension is a complex task, which involves many different cognitive skills and processes. Van den Broek et al. (2005) discusses the meaning of comprehension, and suggest that comprehension is not a unitary phenomenon rather a family of skills and activities. In order to construct deep and coherent understanding the
readers must construct the representation for the text at hand (Kintsch & Kintsch, 2005). Reading comprehension can be defined as an interaction and fusion between the text information and knowledge activated by the comprehender (McNamara & Kintsch, 1996) and the level of understanding of a passage or text. In this respect, it can mean many things, but always includes a meaningful representation of the textual information in the reader’s mind.

Working memory is implicated in academic performances, including reading comprehension in both children and adults (Swanson, 1994), as reading comprehension requires constant integration of incoming information. It also has been hypothesized to be one of the main determinants of individual differences in reading comprehension (e.g., Baddeley, 2007; Daneman & Carpenter, 1980; Perfetti, 1985). For instance, the reader must deal with the incoming information of a text in different levels: At the word level, processes are necessary to encode printed materials, and associate the recognized word with a meaning. At the sentence level, processes are devoted to make sense of the relations among phrases and clauses. At the text level, the readers need to connect successive propositions in order to create an abstract meaning (Long et al., 1996). In short, while reading the reader must store pragmatic, semantic and syntactic information from the preceding text and use it in disambiguating, parsing, and integrating the subsequent text (Daneman & Carpenter, 1980). Therefore, these processes place a lot of demand on working memory, and leads different levels of performance in different individuals.

Comprehension of language depends critically on verbal component of working memory; however, also the central executive and visuo-spatial sketchpad contribute to
comprehension in crucial sense (Gathercole & Baddeley, 1993). Evidence suggests that language processing in reading is influenced by the ability to store and retrieve visual input. Depending on how effectively this input is managed would cause the individual differences in and different level of reading comprehension (Vellutino, 2003). Surprisingly, most of the ability tests do not measure visuo-spatial working memory component and its contribution, although it is fundamental in reading (Vellutino, 2003; Pickering, & Gathercole, 2001).

Engle and Conway (1998) stated that the working memory, especially central executive function’s capacity is essential to comprehension for several reasons (1) The large numbers of words in a sentence, (2) There is a possible confusion about the meaning of words and phrases, (3) the early created representation of the meaning is misleading than the ultimate meaning, which is intended to be communicated (4) and the syntactic structure of the sentences is multifaceted. Cain, Oakhill, and Bryant (2004) have found out that working memory makes a contribution to reading comprehension in 8 to 11-year-olds after the word reading and verbal ability is controlled for.

Individuals differ in their capacity of attention, knowledge and comprehension processes. Differences in Working Memory Capacity (WMC) are thought, in general, to affect the amount of cognitive resources one has to expend on information processing and storage (Baddeley & Hitch, 1974). WMC measures are strongly related to performance in other complex cognitive tasks, but they are particularly crucial component of reading comprehension according to some investigations (Just & Carpenter, 1992; Kintsch & Van Dijk, 1978). These differences might arise from many resources. Since WM’s capacity is assumed to be limited (Baddeley, 1990) working memory becomes a differencing factor
for readers when they read to comprehend. Prior research shows that as the WM demands of a task increase, people’s performance on the task decreases (Baddeley & Hitch, 1974; Just & Carpenter, 1992). The span score obtained from the various measurement tools, such as WMTB-C, provides an idea about the individual’s capacity of temporarily maintaining the information presented for recall and processing (Pickering & Gatercole, 2001).

The adult readers with high working memory spans are able to perform better on reading comprehension tasks than with readers of less working memory capacity (Baddeley, 2007; Linderholm & Van den Broek, 2002). This has been emphasized both in adult and children readers (Daneman & Carpenter, 1980; Seigneuric & Ehrlich, 2005). For instance, Seigneuric and Ehrlich (2005) investigated the role of working memory capacity in children’s reading comprehension in a longitudinal study. One of their primary questions was to what extent reading comprehension could be predicted from working memory capacity in children. Their results of multiple-regression analyses predicted reading comprehension from working memory, non-word reading, and vocabulary measured at a prior time. They claim that as word recognition becomes automated throughout the early grade levels, working memory becomes an important determinant of reading comprehension. Among the variables theoretically related to text comprehension, namely, vocabulary and working memory capacity, vocabulary appeared to be the first predictor to emerge whereas working memory was a significant predictor from Grade 3 only (Seigneuric & Ehrlich, 2005).

Comprehension of language depends critically on verbal component of working memory, namely phonological loop. Gathrcole et al. (2004) emphasized that as a function
of phonological loop a sizable increase in memory capacity occurs as children grow older. They become more experienced in their rehearsal strategies, and accordingly maintain increasing amount of verbal material in their phonological store. On the other hand, the central executive and visuo-spatial sketchpad is equally important. However, these two components of WM have not been explored as much as phonological loop (Gathercole & Baddeley, 1993; Baddeley, 2000). Executive functions have been associated closely with complex cognitive tasks (Boduroglu et al., 2007).

Following the lead of Baddeley and Hitch (1974), most researchers attempting to measure the capacity of working memory and individual differences studies have chosen a dual task procedure, which requires the subject to maintain data in short-term storage while performing some unrelated cognitive activity. Limits in working memory capacity are reflected as a decrement in performance in either the primary or secondary task (Tire & Pena, 1992). As a result, complex reading tasks that require additional resources are more difficult for readers with low WMCs than for readers with high WMCs. The theory also proposes that WMC is an account for individual differences in reading (Daneman & Carpenter, 1980). As a result, some individuals are more affected by increases in WM demands than others (Daily, Lovett, & Reder, 2001).

So Far, the working memory research has been extensive on the effect of working memory capacity on reading comprehension. However, the past research tells us very little about how comprehension processes change with different level of capacity to hold and manipulate information simultaneously. Kintsch and Kintsch (2005) suggest that the disparity between the performance of fluent readers and the struggles of beginners must be the starting point for how instruction in reading comprehension should be
conceptualized. Understanding how readers construct the meaningful representations, or how they cannot, or how they differ in their process might increase the effectiveness of the instruction methods. May be readers should be explicitly taught the methods and strategies of readers with high working memory capacity. This adoption would not be possible without exploring what differs in both groups cognitive processes.

Reading Purpose and Comprehension

We go though different patterns in our comprehension in different reading situations (van den Broek et al., 2001). One factor likely used in determining the variation in the pattern of comprehension outcomes during reading is the purpose (Walker & Meyer, 1980). It is well established in the research literature that individuals process text depending on various factors, such as background knowledge, and reading purpose. The purposes are defined as cognitive representations of what we would like to have happened in the future (Ford, 1992). The depth and richness of our current level of text processing is claimed to be related to the purpose we develop throughout the process we involve (Schutz & Akron, 1994).

A critical role for reading purpose in the comprehension process is implicated by findings that having a purpose while reading influences recall (e.g., Anderson & Pichert, 1978). In another study, having purposes speeded up the process of retrieval of information from memory (Convey, 1990). Furthermore, readers claim to modify their reading strategies according to reading purpose. Barselou (1983) suggested that people organize their thoughts around the information, skills and strategies needed to move them closer to a particular purpose.
For instance; Lorch, Lorch & Klusewitz (1993) investigated different processing requirements of varying reading situations from readers’ perspective. The college students rated different reading situations as demanding. They found that across the 10 reading types introduced, such as reading to apply, reading for research, reading to self inform etc., readers suggested that they make several consistent distinctions with regard to the purpose for which they read. The most obvious distinction made by readers has already been noted several times: School assigned readings and readings by personal choice are perceived as very different in their cognitive demands. School reading was perceived to involve less anticipation of future text events, more attempts at integration, and more rereading. In contrast, entertainment reading was perceived to involve increased effort in finding relations among ideas and events in the text, more anticipation of forthcoming text events, more interest, and more analysis of writing style. Lorch et al. (1993) provided description of text types and reader perception of the text types’ demands.

Narvaez, Van Den Broek, and Ruiz (1999) investigated the effect of reading purpose on inference generation and comprehension behavior in adult readers. They differentiated the out-loud protocols as online comprehension behavior to examine if particular type of inferences or activation is made by the readers. The mean age for the participants was 23 and twenty participants were involved, and they were asked to imagine that they are either reading to study or reading for pleasure. The first relevant result the reading purpose influenced was the pattern of inferences produced by the readers. The study condition led to more evaluations and text repetitions when the participants read an expository text compare to the entertainment condition (Narvaez,
Van Den Broek, & Ruiz, 1999). The interesting aspect of this study was they measured the time course of processing, and it did not differ significantly across the reading purpose condition. Although this might be due to a variety of reasons (Lorch & van den Broek, 1997), it should be emphasized that the effect of working memory capacity on this matter remains unclear.

Linderholm and van Den Broek (2002) investigated how college students with different working memory capacities adjust their cognitive processes when they were given a reading purpose. The researchers used reading span devised by Daneman and Carpenter (1980) to differentiate the capacity of WM as high and low in college level readers. For entertainment they asked the participants to imagine that they are browsing through a magazine and the text caught their attention. Whereas, in study condition, the participants were asked to imagine that they are studying for an essay exam in the class. In the verbal protocol task, the participants read paragraphs, and verbally reflected on how they understand the text and the verbal protocol outcomes were categorized into (associations, evaluative comments, connecting inferences, elaborative inferences, predictive inferences, reinstatement inferences, metacognitive comments, paraphrases, and text repetitions). The researchers found out that all readers adjusted processing to fit the reading purpose; however, when reading for study, low-WMC readers emphasized processes that were the least demanding on their resources but not necessarily beneficial for recall. This illustrates that when a small aspect of reading is manipulated, reading comprehension significantly changes for adults. However, as it is indicated before, the usage of single measure of verbal component of working memory and division of working memory capacity accordingly is one of the handicaps of this study. Furthermore,
WM and reading purpose has been examined in adult’s reading comprehension in this study, the interaction effect of the two factors among those have not been explored in children yet. Additionally, the division of working memory capacity by using different measures needs to be investigated.

Van den Broek et al. (2001) also examined the relationship between the readers’ purposes and how much inference generation occurs when they read expository text as well as frequency of recall for the text. 82 college students involved in this study and read texts for the purpose of either study or entertainment. On-line inference generation was recorded by think-aloud procedures, and off-line memory was assessed by free recall. Notion of Standards of Coherence provided a framework for their exploration of how reading purposes affect the outcomes of the reading comprehension (Van den Broek et al., 2001). Standards determine when adequate coherence is attained and when additional information is needed from prior learning, or from background knowledge (Lorch, Lorch & Klusewitz, 1993). According to this view, as the reader move forward though the text, in order to maintain the coherence in their understanding, they employ many inferential activities as they read, especially causal and referential. These were chosen as standards of coherence because these apply across reading situations and individuals (van den Broek et al. 2004). They found out that reading purpose strongly influences the cognitive processes of inference generation: Readers with a study purpose produced more coherence building inferences, whereas readers with an entertainment purpose generated more associations and evaluations. The results indicate that inference generation during reading is partly strategic and is influenced systematically by reading purpose.
The researchers suggest that reading purposes influence readers’ *standards of coherence*, which in turn influence the types of inferences and final memory products that they produce.

The methodologies in these three studies (Linderholm & van den Broek, 2002; Van den Broek et al., 2001; Narvaez et al., 1999) were similar: Adult readers were involved and they were asked to imagine that they are either reading for study or reading for pleasure. The think out loud protocols are considered the best way to investigate the reading comprehension process (Ericsson & Simon, 1993).

As indicated the evidence supports that readers can adjust their processes (Linderholm, & Van Den Broek, 2002; Lorch, Lorch, & Klusewitz, 1993) according to distinct contexts. Reader with a variety of well developed processing strategies will not be an effective reader if he or she fails to recognize that particular strategies are relevant in the context of reading is relevant (Lorch, Lorch, & Klusewitz, 1993). In this respect the dominant processes in young readers’ comprehension, especially when the difference is considered in their working memory capacity and they are given purpose for reading still remains unclear, as there is no research with young children on the issue. Consequently, there is no previous work has assessed reading comprehension in young readers aged 11 to 12 years old, and how their processing of the text differs when a particular reading purpose is provided with combination of traditional methods of reading comprehension, for instance reading comprehension questions and think out loud protocols. Furthermore, we know that adult readers adjust their processes to maintain the coherence; however this is not intensely investigated in young readers when external factors, such as given reading purposes affect comprehension when WM is a determinant.
Aims of the Current Study

In educational settings the goal of reading is usually to learn from a text and to construct representations of the content in the text. This introduced information will be remembered and can be used effectively when it is needed later. In order to construct coherent representations, it could be beneficial to direct the goals of reading. Still little is known about what would be the cognitive processes in young readers produced by different given reading purposes. Another question of interest is how we can provide young readers the tools, so that they can construct representations to be used in the future. Therefore, this study was critical to explore what kind of cognitive processes young readers are involved when they read with different goals. Providing the reading purpose may improve the reading comprehension performance for young readers with different working memory capacities.

The results of this study will have both practical and theoretical implications: Differentiated reading purposes might improve the reading comprehension instruction in educational settings. Moreover, as indicated above, the existing literature is intensive about the adult performance in the similar research set ups. However, limited number of research has explored the performance of young readers. Thus, the findings intends to bridge the gap between the adult literature and children’s as well as applied and laboratory contexts. Given that the comprehension is dynamic process depending on the unique factors of the context, and individual differences in readers; one of the key goals of the present study is to investigate the role of working memory capacity and reading purpose on the cognitive processes involved in reading comprehension of young readers. It is predicted that the one of the reading purpose conditions (reading to take a
test or reading to play a game) will prove to be a favorable context for young readers with different working memory capacity.

**Research Hypothesis and Research Questions**

The main hypothesis in this study is working memory capacity (High WMC or low WMC) and reading goal (reading for taking a test or reading for playing a game) will elicit different comprehension processes in young readers. In addition, I predict that the readers’ frequency of text recall will differ according to the reading purpose condition (test versus game) and the WMC (high versus low). Another assumption in this study is that the performance on comprehension and application measure’s outcomes will differ according to WMC and reading purpose provided.

The three primary research questions raised in the current study are

1) Do working memory capacity and reading purpose influence text comprehension process in young readers?
2) Do the comprehension and application outcomes differ depending on the WMC and reading purpose provided?
3) Does free recall frequency vary as a function of reading purpose and WMC?

**Method**

**Participants**

Initially there were 43 participants; elementary school grade five and six students participated in this study. However 4 of the types could not be transcribed due to inaudible voices in recordings. Eventually 23 female and 16 male participants, 39 in total, were included in this study. The participants were chosen from Montreal public schools, in which the main teaching language is English. Montreal is a multilingual city; therefore,
the participants reflected the diverse ethnic backgrounds and languages spoken at home. The parental background questionnaires and children’s self ratings revealed that 31% of the participants’ first language was English, and 23% of the total participants’ first language was French. The remainder of the participants, 46% considered both languages as their first language.

The age range of the students was between 10 years, 5 months and 12 years, 5 months. The mean age was 11 years and 9 months (SD=5.87, in months). After the initial school selection, the students were randomly assigned to one of the reading purpose conditions (test or game). We asked all of the participants, if they have ever played the game Go, in order to control the influence of background and knowledge on the game. Furthermore, the sessions were monitored to ensure if the participants knew the game before. There was only one participant who played the game; his data was discarded and removed from the analyses. Parental Background questionnaires were also used to explore, if the participants have any learning problems. Participants that were identified to have any type of learning problems were excluded from the study.

Materials

Text. All participants in this study read the same text titled “How to play the ancient, oriental game of GO: The Way to GO” by Karl Baker (1998). The author’s consent was sought for using the text in the study. The text was edited both by a go game player, and the author in order to ensure the suitability of the text level for the targeted audience. The text was 755 words long and adapted from the original. It was presented to the participants on an 8 ½ X 11 size page, double spaced, and on two separate pages.
This particular text was chosen to minimize the participants’ ability to involve their prior knowledge of any game. In addition, the task created for this study could be identically used for dual purpose, test or game.

After the pilot testing, a dot was placed at the end of each paragraph, indicating that the children have to stop and reflect their understanding of the paragraph they just read. The aim of this addition was to indicate the benchmarks for all the participants, where they need to stop and talk about their understanding.

There were two illustrations in the text. The first one was two Go players from Far East, playing the game, and the second one was a Go board showing grids from a distance. The number of illustrations was limited to avoid providing visual aid, which could contribute to the students’ comprehension. The entire text can be found in Appendix A.

An audio tape recorder was also used for the verbal protocol task, to record the participants’ verbal responses for later transcription, coding and analyses.

A Go game board grid was generated by a computer for each participant. 9 by 9 lines were drawn on A4 sized paper and the grid was used for the application part of this study.

Black markers also used for generating the black stones. For the purpose of recording each participant’s performance markers were more appropriate than real stones.

Measures

Students were assessed on a range of working memory and comprehension skill measures according to success in reading comprehension. These measures are reported below.
Aptitude and Achievement Measures

Peabody Picture Vocabulary Test (PPVT-III). PPVT-III is devised to assess an examinee’s receptive (hearing) vocabulary for Standard English. It functions as a screening test of verbal ability, or as an element of a comprehensive test of cognitive processes. Vocabulary is a fundamental criterion in reading comprehension (Vellutino, 2003); therefore this measure was included to assess the examinee’s vocabulary level. PPVT-III took 10-12 minutes in average for each examinee (Dunn & Dunn, 1997). In occasional cases, “reverse rule” was applied: If the examinee fails to identify any of the correct answers at the starting level, the preceding block of testing was applied until the examinee was able to successfully complete one block.

Woodcock-Johnson-III Passage Comprehension. This section of the Woodcock-Johnson-III measure consists of cloze procedure. The participant has to read a single paragraph and fill in the blank at the end of each paragraph with the most suitable word. Each paragraph was presented to the examiner on a different page. Administration of this measure takes 10-15 minutes depending on the examinee’s capacity to proceed.

Working Memory Test Battery for Children (WMTB-C). This is a standardized test to measure the performance and capacity of different components of working memory (Pickering & Gathercole, 2001) to provide comprehensive assessment of the three component structure of the Working Memory Model proposed by Baddeley and Hitch. WMTB-C provides the assessment of 5 to 15 year olds.

The selected subtests were applied in a sequence, which was proposed by the WMTB-C manual followed in this study: (1) Word List Recall; (2) Non-Word List Recall; (3) Listening recall; (4) Block Recall and (5) Counting recall subtests were
applied. This sequence is suggested to avoid overtaxing one component of working memory in consecutive applications (Pickering & Gathercole, 2001). The only exception was the counting recall task, which was applied separately, in the second testing session. In occasional cases, reverse rule was applied: If the examinee could not obtain 4 correct answers in the starting level, the preceding block of testing was applied until the child was able to successfully complete one span.

Five different subtests of WMTB-C were administered:

1. In Word List Recall (a measure of phonological loop function) subtest, there are seven spans, in each span there are six trials of items. The first span has 1 word to recall, and the seventh span has 7 words. In other words, the number of words increases across spans. The examiner reads aloud a list of one-syllable words and the examinee was expected to accurately recall both the content and the sequence. Each trial consists of the same number of words in a particular span to recall. When the examinee correctly responds to 4 trials in a particular span, this gives the signal to the examiner to continue with the next level. The discontinuation rule is 3 wrong trials in any span. The number of correct trials sums up to a total number of correct trials. When the examinee discontinued due to 4 wrong trials in any set, the last successive span was considered as the child’s working memory span for that particular subtest. Test–retest reliability coefficient is .64 for word list recall, for children aged 11-12 years (Pickering & Gathercole, 2001).

2. Nonword list recall (a measure of phonological loop function) subtest was given to the participants. In this subtest, non-words (6 sets of 6) are presented to the child in spoken form using an audiocassette recorder. The examinee is required to repeat each item once it was presented. In this subtest, there are six spans, in each span there are six
trials of items. The first span has 1 non-word to recall, and the sixth span has 6 non-words. In other words, the number of words increases across spans. The examiner reads aloud a list of non-words and the examinee is expected to accurately repeat both the content and the sequence. Each trial consists of the same number of non-words in a particular span to recall. The repetition attempt is scored as correct if there is no phonological deviation from the target form. Once the examinee responded correctly to 4 trials in a particular span, the examiner continued with the next level. The cut-off for discontinuation rule was 3 wrong trials in any span. The number of trials correct sums up to a total number of trials correct. When the child discontinued due to 4 wrong trials in any set, the last successive span was considered as the child’s working memory span for that particular subtest. Test–retest reliability coefficient is .43 for non-word list recall, for children aged 11-12 years (Pickering & Gathercole, 2001).

(3) In Listening Recall (a measure of central executive function) subtest, on each trial the examinee hears a sequence of simple sentences to which they had to response ‘true’ or ‘false’. Later, they are asked to recall the final word of each sentence in the sequence in which the sentences were presented. In this subtest, there are six spans, in each span there are six trials of items, consisting of different number of sentences. The first span has 1 sentence where the child has to identify if the sentence is true or false and recall the last word, and the last span have 6 sentences. The examiner reads out loud the list of sentences and the examinee decides if the sentence is true or false. Next the examinee is expected to recall the last word of the each sentence. Each trial consists of same number of sentences in a particular span to recall. The repetition attempt is scored as correct if there is no phonological deviation from the target form. When the examinee
correctly responds to 4 trials in a particular span, this gives the signal to the examiner to continue with the next level. The discontinuation rule is 3 wrong trials in any span. The number of correct trial sums up to a total number of correct trials. When the child discontinued due to 4 wrong trials in any set, the last successive span was considered as the child’s working memory span for that particular subtest. Test–retest reliability coefficient is .38 for listening recall, for children aged 11-12 years (Pickering & Gathercole, 2001).

(4) Counting Recall (a measure of central executive function) subtest was introduced as a distracter task in between reading the text and free recall of the text. The counting recall test was also taken from the WMTB-C. In this subtest, the child is required to count the number of dots in an array, and then recall the total number of dots in the arrays that were presented. A booklet is placed in front of the examinee, consisting of pages displaying three, four, five, or six red dots in a box. The test trial begins with one page of a dot array, and increases by one page in each level, until the child is unable to correctly recall four trials. The number of correct trials was scored for each child. Test–retest reliability coefficient is .48 for counting recall, for children aged 11-12 years (Pickering & Gathercole, 2001).

(5) During the Block Recall (a measure of visuo-spatial sketchpad function) subtest the experimenter taps a sequence of blocks arranged unsystematically on a three-dimensional board. The examinee is required to tap on the each block sequence exactly in the same order, which is presented by the experimenter to reproduce the sequence in the same order. In this subtest, there are nine spans, in each span there are six trials of items. The first span has 1 block to recall, and the ninth span has 9 blocks. In other words, the
number of words increases across spans. The examinee is expected to accurately repeat the sequence. Each trial consists of same number of blocks in a particular span to recall. When the examinee correctly responds to 4 trials in a particular span, this signifies the examiner to continue with the next level. The discontinuation rule is 3 wrong trials in any span. The number of correct trials sums up to a total number of correct trials. When the child discontinued due to 4 wrong trials in any set, the last successive span was considered as the child’s working memory span for that particular subtest. Test–retest reliability coefficient is .43 for block recall, for children aged 11-12 years (Pickering & Gathercole, 2001).

**Comprehension-Application Measures**

In this study different tasks were developed in order to measure the outcomes of reading. These were (1) visual comprehension question problems of the game Go, (2) the verbal comprehension questions of the game Go, and (3) the application of the game. These measurement tools intended to include the content of the text and relevant material. The text was removed from the testing scene as soon as the participants finished reading. During testing, the participants could not go back and refer to the text. Moreover, the participants were not provided any information, or explanation. These tasks were piloted to ensure suitability, reliability and validity. After the pilot testing the tasks were re-adjusted.

**Visual Comprehension Questions of the Game.** The visual comprehension problems were taken from the same text by Baker (1998). In order to ensure the suitability of the questions’ difficulty levels for the target audience and for the purpose of this study, the questions were simplified by a Go game player, and the researcher in this study. The
purpose of the task was to measure comprehension level of readers and how much they could apply the rules to a real context.

During this task, the participant was given 12 Go problems accompanied by both visual configurations of the question verbal questions (see appendix B for examples of this task). The first two were practice items. Each participant received the same explanation and feedback, regardless of the answer they provided. This was to control confounding factors of different instruction. The participants received no feedback on their performance after the two practice trials.

The questions were presented one at a time on an A4 size white paper. The experimenter asked the participant to read the questions silently and provide the best answer they could think of. The participants could read the questions as many times as they needed to understand. The participants could not go back and refer to the text, if they were unsure of their answers. They were not provided any information, or explanation, if they asked any question. The text was removed from the testing scene right after they finished reading.

*Verbal Comprehension Questions.* Reading comprehension was also measured by a combination of different types of questions related to the text. A balance of multiple choice, cloze procedure, and “true or false” questions were used in this task. Similar types of questions were chunked together. Some of the questions were taken directly from the text, while others were created on the bases of possible inferences from the text. The experimenter asked the participant to read the questions silently and provide the best answer they can think of. The questions can be found in Appendix C.
Application of the Game Go. In order to assess how the participants would apply the rules of the game, right after the comprehension application questions participants played the game with the examiner. Each participant received a black marker to use as black stones. The total amount of time for the game differed between 3 minutes to 4 minutes, depending on the participant’s performance. The examiner rated the play based on participants’ performance on some of the observable rules of the game; such as placing stones on the points rather than on the squares, capturing the examiner’s stones, protecting their own territory etc.

Procedure

After the initial school and classroom selections, the participants were randomly assigned to either the test or the game condition. They had unlimited amount of time for reading the text, and summarizing their understanding. Pilot testing of the materials revealed that long sessions might cause fatigue and distorted performance. Therefore, the total time of testing was divided into two sessions. The first and second testing dates were a maximum of 30 days apart. Two trained researchers from McGill University worked with the participants, in a quiet room in a school setting. The workspace was close to the participants’ classrooms. Each testing session took 30 to 40 minutes per participant, depending on the child’s ability to proceed. Upon completion of the second and final session, the participants were informed that the condition they involved had a dual purpose of both game and test. In the first session, Woodcock-Johnson-III Passage Comprehension, and Working Memory Test Battery for Children (WMTB-C) subtests; in the second testing session, Peabody Picture Vocabulary Test (PPVT-III) and non-standardized of Comprehension-application measures (Verbal protocol task, visual
comprehension questions, verbal comprehension questions, and the application of the game) were applied. The WMTB-C counting recall subtest was administered in between the verbal protocol task and free recall task.

*Verbal Protocol Task.* A verbal protocol task was administered. This task included an expository text. The participants were asked to read the text silently, one paragraph at a time. However during the session, if the participant chose to read out loud, in spite of the reminder at the beginning, the researcher did not intervene. All responses were recorded by a tape-recorded and were transcribed later for analysis. The participants were acknowledged about the type recorder, and a verbal consent was sought to ensure that the participant would be comfortable with his/her voice recorded.

The “test purpose” group individuals were told: “I am interested in how young people read. Today you are going to take a test after you read a text. The purpose of this particular reading is to take the test afterwards. First, I need you to read the text silently. You need to read one sentence at a time. When you start reading it, after each paragraph you will see a little dot. This indicates that you should stop and talk about what you understand from what you have just read. When you finish reading the whole piece, I will give you different tests to complete. It is important that you understand what is written in the text, because all the tests will be related to what you read here. Remember to read carefully, so that you will do well in the test.’’

The “game purpose” group individuals were told: “I am interested in how young people read. Today you are going to play a game after you read a text. The purpose of this particular reading is to play a game afterwards. ‘First, I need you to read the text silently. You need to read one sentence at a time. When you start reading it, after each
paragraph you will see a little dot. This indicates that you should stop and talk about what you understand from what you have just read. When you finish reading the whole piece, I will give you different games to complete. It is important that you understand what is written in the text, because all the games will be related to what you read here. Remember to read carefully, so that you will do well in the game.”

When reading for playing a game, students were told that after reading the text they would be asked to play a game, when reading for taking a test, the students were told that after reading the text, they would be asked to take a test.

Distracter Task. Working Memory Test Battery for Children (WMTB-C) counting recall subtest was used in between the verbal protocol and the free recall task as a distractor task. The reader’s recall of their understanding of the text was delayed by having them complete a distracter task prior to the recall. WMTB-C counting recall task was chosen particularly. Because it is short and does not completely rely on verbal comprehension skills. The verbal protocol task consisted of reading a text and talking about understanding of the text before the counting recall subtest of WMTB-C were both verbal tasks. The counting recall subtest would be less demanding on phonological loop and central executive. This particular sequence (reading the text-verbal protocol task, counting recall task, and then free recall task) was to avoid overtaxing one component of working memory in consecutive application (Pickering & Gathercole, 2001).

Free Recall Task. Upon completion of the distracter task, WMTB-C counting recall subtest, the reader was asked to talk about their understanding of the text. Participant’s recall was tape-recorded. The participants were asked to talk about whatever they would like to tell related to the text.
Visual Comprehension Questions of the Game. In this task the participant was given 12 Go configuration problems accompanied with verbal questions (Appendix B). The first two items were for practice trials. Each participant received the same explanation and feedback, even though they provided the right answer. This was to control confounding factors of different instruction. The participants received no feedback on their performance after the two practice trials. The questions were presented one at a time on an 8 ½ X 11 size white paper. The experimenter asked the participant to read the questions silently and provide the best answer. The participants could read the questions as many times as to comprehend. The participants could not go back and refer to the text, if they were unsure of their answers. They were not provided any information, or explanation, if they asked any questions. The text was removed from testing scene right after they complete reading.

Verbal Comprehension Questions of the Game. The experimenter asked the participant to read the questions silently and provide the best answer they can think of. The participants could not go back and refer to the text, if they were unsure of their answers. They were not provided any information, or explanation, if they asked any question. The text was removed from testing scene right after they completed reading.

Application of the Game. Each participant played the game with the examiner, who has the knowledge of how to play the game. Total amount of time for the play differed between 3 to 4 minutes, depending on the participant’s performance. The examiner rated the play based on participants’ performance on some of the observable rules of the game; such as placing the stones on the points rather than on the squares, capturing the examiner’s stones, protecting their own territory etc. In order to assess how the
participants would apply the rules of the game, each participant was provided a black marker. This marker was utilized as black stones. When all the children completed their final sessions, they were acknowledged that the tasks they involved had dual purpose of both game and test.

Data Scoring

Verbal Protocol and Free-Recall Task

The verbal protocols of the participants were transcribed from audiotapes for analysis. Each verbal protocol response was divided into complete idea units. Idea units were determined by discussion between the two independent raters until 90% agreement was attained. Instead of the sentences uttered, it was appropriate to use idea units, because of two reasons: (1) In some of the cases, the participant did not complete the sentences, but still expressed ideas from the paragraph or the text; (2) in other cases, within one sentence there were many idea units combined from the text. Therefore, looking at the frequency of idea units would be an accurate measure of the comprehension processes.

All audio-tapes were transcribed by the researcher. The transcriptions were scored by two researcher assistants from McGill University. The primary rater was a research assistant from the linguistics department of the same university.

In order to negotiate mutual understanding of the categories, a discussion session was carried out, including the two raters and a panel of judges from the same research laboratory. The disagreements were resolved by negotiating new categories, or creating a collection of common themes in each category. One rater scored the remainder of the verbal protocols. 25% of the data was scored separately by two different raters. The raters
were blind to the groups, to which participants were allocated. The inter-rater reliability coefficient was 87% for the scoring. Each idea unit both from verbal protocol task and free recall was categorized into one of the seven cognitive processing categories: 

*Associations* were the retrieval of information not related to text coherence. Idea units were accepted as associations when the reader referred to concepts or ideas outside of the text. It refers to concepts that were brought to mind by the text that were not intended to enhance coherence (e.g., “This game Go, reminds me of a game we play with my brother, called capture”); *Evaluative comments* were opinions about the text, which refer to participants’ opinion about the text (e.g., “This game is really fun.”). When readers attempt to explain the contents of the current sentence on the basis of background knowledge, or outside resources, these inferences are categorized as *elaborative inferences* (Linderholm & Van den Broek, 2002). For instance, in the text or the game rules there is no taking over rule, however the reader comments on the paragraph basing on their previous knowledge of invasion games. (e.g., “We have to take over the partner’s territory.”). *Predictive inferences* were idea units, which entail anticipating what will occur next in the text. Predictive inferences were forward inferences that anticipate upcoming text. We also included the readers predictions of how the game might look like when or while they play (e.g., “In order to capture your partner’s stones, you will have to look for the weak group.”). *Connecting-reinstatement inferences* involved explaining the contents of the current paragraph by connecting its meaning with the preceding sentences or any information from the text. When readers attempt to provide an explanation for the current sentence on the basis of prior text information that was not in the immediately preceding sentence, these processes were thought of as connecting-reinstatement
inferences (e.g., “And, put their stones into prison, you try to make attached groups together that are called units.”). This category was created by combining two distinct categories in Linderholm and van den Broek’s study (2002). *Metacognitive comments* occurred when readers reflect on their understanding or lack of understanding of text information (e.g., “I understood the paragraph, but I did not get some meanings!”).

*Paraphrases* occurred when readers put the current sentence into their own words. Paraphrases were the comments that capture the gist meaning of a sentence (e.g., “So, The game go was created in China.”).

Initially, the previous work of Linderholm and van den Broek (2002) scoring categories would be replicated in our entire scoring schemata. In the Linderholm and van den Broek (2002) study the researchers asked the participants to talk about their understanding after every sentence. However, in our study we asked the participants to reflect their understanding after every paragraph. Some of their categories, such as connecting inference and re-instatement interference, were not suitable for the present study. When a participant commented they talked about the paragraph as a whole not sentence by sentence. In this case, the connecting inference and re-instatement interference would not be differentiated as the paragraph is read as a whole, rather than separate sentences one at a time. Thus, our new category reflected the combination these two.

*The Comprehension-Application Tasks*

The correct answers of visual and verbal comprehension questions were pre-established. For the application part of the game the participants were rated on the basis of pre-established criteria (such as black started first, the participant placed the stone on
points, did not try to move the stones, grouping the stones to create territory etc.). For each right answer the participants received score of 1, and 0 for wrong answers.

Results

Preliminary Analyses

All analyses involving standardized cognitive ability measures, and comprehension application measures were based upon $N = 39$ observations. The total number of participants in each condition was 21 for the test purpose and 18 for the game purpose.

Following the methodology and reasoning of Linderholm and van den Broek (2002) study, in order to identify the participants working memory capacity, first, the five subtest standard scores (Word List Recall, Nonword List Recall, Listening Recall, Counting Recall, Block Recall) of Working Memory Test Battery for Children (WMTB-C) were combined to create a new composite score. In the original study of Linderholm and van den Broek (2002) a reading span task was used to differentiate the WMC. They recorded the total number of words accurately recalled on this reading span task, and the upper and lower third of the frequency distribution of total words recalled by each participant was used to decide low- from high-WMC readers. Due to the small sample size, discarding the data of subjects was not considered in this study. Therefore, median split was the closest approach to the original study. We did a median split into two groups based on the new pooled scores of WMTB-C subtests. The upper half of the median score was accepted as high working memory capacity and lower part was accepted as low working memory capacity. The number of high and low WMC readers differed in each condition: In the test condition there were 12
high WMC and 9 low WMC readers; whereas in the game condition, there were 9 high WMC and 9 low WMC readers. In total, there were 18 participants in game condition, 21 in test condition.

The vocabulary age equivalence mean was 11.7 ($SD = 3.1$, in years), passage comprehension age equivalence mean score was 9.10 ($SD = 1.18$, in years). The mean score for global working memory capacity was 120.90 ($SD = 8.79$). These mean scores indicates that the sample of this study was above average performance on working memory test battery global working memory scores. For standardized measures all analyses were based on standard scores. Although vocabulary level measured by PPVT-III and passage comprehension measured by WJ-Passage Comprehension subtest were initially thought to be used as covariates in all the analyses to be reported, however no significant main effects or interactions were associated with them. Therefore, for economy of presentation the results of the analyses run are not presented here. The means and standard deviations of all standardized measures are presented in Table 1. Preliminary inspection of the data revealed no signs of marked kurtosis or skew in the data for any working memory, vocabulary or passage comprehension standard scores. The mean scores for standardized measures can be found in Table 1, in Appendix D.

One of the assumptions of ANOVA is that the distributions in each of the groups are normal. In order to meet the assumptions of ANOVA analyses, the outcome data was explored, if it is normally distributed for the sample involved in this study. Non-normality is commonly characterized in terms of two parameters, namely, skewness and kurtosis. These are two indicators, if the data is normally distributed. The departure of the
distribution from normality, and the distribution had to be normalized for some outcome measures before we preceded with appropriate analyses methods (Tabachnick & Fidell, 2001). Therefore, some of the outcome measures in cognitive processing categories (1) Verbal protocol task: Associations and connecting-reinstatement inference; (2) Free recall task: Evaluative comments and predictive inference were adjusted to meet the assumption of normality. The adjustment was done in the statistical program SPSS, by transforming and computing the values of above mentioned categories by square root function (SQRT).

Another rater, who was blind to the treatment conditions of the participants scored 25% of the out loud protocol and free recall data. The interrater reliability was (r=.87), which was found to be suitable.

Partial $\eta^2$ values were reported for each analysis, even if the results were not significant at $p < .05$. According to Stevens (2002) exploring effect sizes is one way to obtain practical significance. Probably, due to the small sample size obtained in this study, many results were not statistically significant; the effect sizes might be useful to identify practical significance and given direction for future research.

According to Stevens (2002) exploring the effect sizes is one way to obtain practical significance. Effect size is simply a way of quantifying the size of the difference between two groups. It is a way of quantifying the difference between two groups that emphasizes the size of the difference rather than confounding this with sample size (Cohen, 2001). In usual cases, it is a determinant to explore whether a statistically significant difference is a difference of practical concern. It is important to know the statistical significance of a result, since without it there is a danger of drawing firm conclusions from studies where
the sample is too small to justify such confidence. However, statistical significance does not tell you the most important thing: the size of the effect. Moreover, it does not place the emphasis on the most important aspect of an intervention – the size of the effect – rather than its statistical significance. In practical situations, effect sizes are helpful for making decisions (Coe, 2002).

Although nearly in all of the case the results were not statistically significant, we still proceeded with reporting the effect sizes. Medium to large effect sizes were reported to illustrate the difference between group means. It should be noted that the effect sizes reported here do not have intention to demonstrate what we have not found in our research, significant results. Instead, we are looking to see of effect size to be functional to identify practical significance and provide direction for future research. estimates can be practical, and can direct next research hypothesis for a future research agenda. Since the effect sizes can serve as an important mechanism for communicating our professional knowledge and tendencies. Still, the results should be cautiously interpreted as they are not significant statistically.

Main Analysis

The data are presented in four sections: (a) The effect of reading purpose and WMC on adjusting cognitive processes during reading, (b) The effect of reading purpose and WMC on recall, (c) The effect of reading purpose and WMC on adjusting cognitive processes during the free recall task, and (d) The effect of reading purpose and WMC on performance in the Application-Comprehension Measures.

(a) The effect of Reading Purpose and WMC on Adjusting Cognitive Processes during Reading
Our first hypothesis was that cognitive processes, measured by verbal protocol responses, would vary as a function of reading purpose and WMC in young readers. The independent variables were reading goal and the working memory capacity scores and the dependent variables were seven categories of cognitive process responses uttered by the readers (associations, evaluative comments, connecting-reinstatement inferences, elaborative inferences, predictive inferences, metacognitive comments, and paraphrases). The main question explored in this study concerned the relationship between the effect of reading purpose and working memory capacity on adjustment of cognitive processes while reading text. Separate analyses of variance (ANOVAs) between participants were applied for each verbal protocol response category separately. The aim was to detect differences between groups (test versus game condition or high WMC versus low WMC) or interaction effect in the frequency of idea units in each separate cognitive process response category. ANOVA was the appropriate analyses because we intended to investigate the effect of two independent variables on outcome measures concurrently, in addition the past research used the same analyses for similar designs (Linderholm & van den Broek, 2002). The independent variables of interest were WMC (low vs. high) and reading purpose (test vs. game), and the dependent variables were cognitive processes (associations, evaluative comments, connecting-reinstatement inferences, elaborative inferences, predictive inferences, metacognitive comments, and paraphrases). The number in each category represents the frequency of idea units uttered by the participant. $F$ statistics that reached an alpha level of .05 were considered significant and are reported. An alpha level of .05 was used for all the statistical tests.
The hypothesis that readers in test versus game groups or high versus low WMC was not confirmed by analyses of variance (ANOVAs) for each response category. There were no significant differences found between the groups $p>.05$. There was no main effect difference for either condition or for working memory capacity measured by WMTB-C in this category of responses between groups. These data, presented in Table 2 clearly shows the pattern of mean scores and standard deviations across groups for each cognitive process response category separately.

**Associations.** A two way ANOVA revealed that utterance of associations did not differ across condition groups ($F(1,34) = .277$, ns; $\text{Eta}^2 = .008$), or WMC ($F(1,34) = 1.342$, ns; $\text{Eta}^2 = .038$). The interaction effect was also not significant ($F = .018$, ns; $\text{Eta}^2 = .001$).

**Evaluative Comments.** A two way ANOVA revealed that utterance of evaluative comments did not differ across condition groups ($F(1,34) = .328$, ns; $\text{Eta}^2 = .010$), or WMC ($F(1,34) = 1.112$, ns; $\text{Eta}^2 = .032$). The interaction effect was also not significant ($F(1,34) = .567$, ns; $\text{Eta}^2 = .016$).

**Elaborative Inference.** A two way ANOVA revealed that utterance of elaborative inference did not differ across condition groups ($F(1,34) = .81$, ns; $\text{Eta}^2 = .002$), or WMC ($F(1,34) = .857$, ns; $\text{Eta}^2 = .025$). The interaction effect was also not significant ($F(1,34) = .218$, ns; $\text{Eta}^2 = .001$).

**Predictive Inference.** A two way ANOVA revealed that utterance of predictive inference did not differ across condition groups ($F(1,34) = .258$, ns; $\text{Eta}^2 = .008$), or WMC ($F(1,34) = 2.425$, ns; $\text{Eta}^2 = .067$). However, the WMC partial $\text{Eta}^2$ indicates a medium effect size. Readers with low working memory capacity expressed more predictive inference
(M=3.76, SD=2.93) than the readers with high working memory capacity (M=2.48, SD=2.16). The interaction effect was also not significant ($F(1,34)= .329$, ns; $Eta^2=.010$).

Connecting-Reinstatement Inference. A two way ANOVA revealed that utterance of connecting-reinstatement inference did not differ across condition groups ($F(1,34)= 2.573$, ns; $Eta^2=.070$), or WMC ($F(1,34)= 2.573$, ns; $Eta^2=.070$). However, both the condition and WMC partial $Eta^2$ indicate a medium effect size. In game condition, readers produced more connecting reinstatement inferences (M=3.47, SD=3.12) than test condition (M=2.10, SD=2.41). The readers with low WMC produced more connecting-reinstatement inferences (M=3.47, SD=3.45) than readers with high WMC (M=2.10, SD=2.02). The interaction effect was also not significant ($F(1,34)= 1.282$, ns; $Eta^2=.036$). The medium effect size were obtained before and after the data was adjusted to be normally distributed.

Metacognitive Comments. A two way ANOVA revealed that utterance of metacognitive comments did not differ across condition groups ($F(1,34)= .389$, ns; $Eta^2 =.011$), or WMC ($F(1,34)= .024$, ns; $Eta^2=.001$). The interaction effect was also not significant ($F(1,34)= 2.192$, ns; $Eta^2=.061$). However, the interaction between reading goal condition and WMC partial $Eta^2$ indicates a medium effect size. The students with high WMC in game condition (M=3.67, SD=2.35) uttered more metacognitive comments than in other groups.

Paraphrase. A two way ANOVA revealed that paraphrasing frequency did not differ across condition groups ($F(1,34)= .788$, ns; $Eta^2 =.023$), or WMC groups ($F(1,34)= .150$, ns; $Eta^2=.004$). The interaction effect was also not significant ($F(1,34)= .689$, ns; $Eta^2=.020$).
(b) The effect of Reading Purpose and WMC on Adjusting Cognitive Processes during the Free Recall Task

The next step in the analyses was to explore if the cognitive processes responses differed as a function of reading purpose and WMC in free recall task. The mean performance scores across groups for the analyses of free recall protocol task were presented in Table 3.

Associations. A two way ANOVA revealed that utterance of associations in free recall task did not differ across condition groups \((F(1,34) = .584, \text{ ns}; \text{Eta}^2 = .015)\), or WMC \((F(1,34) = .110, \text{ ns}; \text{Eta}^2 = .003)\). The interaction effect was also not significant \((F(1,34) = 1.276, \text{ ns}; \text{Eta}^2 = .036)\).

Evaluative Comments. A two way ANOVA revealed that utterance of evaluative comments differed across reading goal groups \((F(1,34) = 4.505, p<.04; \text{Eta}^2 = .117)\). The readers in game condition \((M=.41, \text{ SD}=.75)\) made more evaluative comments than test condition \((M=.26, \text{ SD}=.21)\). However, the frequency of evaluative comments were similar across WMC groups \((F(1,34) = 1.485, \text{ ns}; \text{Eta}^2 = .042)\). No interaction effect was detected \((F(1,34) = .336, \text{ ns}; \text{Eta}^2 = .010)\). This significant result was obtained after the evaluative comments outcome data was adjusted to be normally distributed.

Elaborative Inference. A two way ANOVA revealed that utterance of elaborative inference did not differ across condition groups \((F(1,34) = .054, \text{ ns}; \text{Eta}^2 = .002)\), or WMC \((F(1,34) = .054, \text{ ns}; \text{Eta}^2 = .002)\). The interaction effect was also not significant \((F=1.085, \text{ ns}; \text{Eta}^2 = .031)\).

Predictive Inference. A two way ANOVA revealed that utterance of predictive inference did not differ across condition groups \((F(1,34) = .107, \text{ ns}; \text{Eta}^2 = .003)\), or WMC \((F(1,34) = .
However, the WMC partial Eta² indicates a medium effect size.

Readers with low working memory capacity expressed more predictive inference (M=1.14, SD=.98) than the readers with high working memory capacity (M=.678, SD=.719). The interaction effect was also not significant ($F(1,34)=.780$, ns; Eta²=.005).

Connecting-Reinstatement Inference. A two way ANOVA revealed that utterance of connecting-reinstatement inference did not differ across condition groups ($F(1,34)=.701$, ns; Eta²=.020), or WMC ($F(1,34)=.068$, ns; Eta²=.002). The interaction effect was also not significant ($F(1,34)=.205$, ns; Eta²=.006).

Metacognitive Comments. A two way ANOVA revealed that utterance of metacognitive comments did not differ across condition groups ($F(1,34)=.440$, ns; Eta²=.018), nor WMC ($F=.0262$, ns; Eta²=.037). The interaction effect was also not significant ($F(1,34)=.066$, ns; Eta²=.096).

Paraphrase. A two way ANOVA revealed that utterance of paraphrase did not differ across condition groups ($F(1,34)=.109$, ns; Eta²=.074), or WMC ($F(1,34)=.988$, ns; Eta²=.000). However, the test condition partial Eta² indicates a medium effect size.

Readers in test condition expressed more paraphrasing (M=10.24, SD=4.85) than the readers in game condition (M=7.88, SD=3.31). The interaction effect was also not significant ($F(1,34)=.996$, ns; Eta²=.020).

\(c)\ The Effect of Reading Purpose and WMC on recall

The main interest in this part of the analyses was to test if the readers in test purpose group would recall more than the game purpose group. Analysis of variance (ANOVA) was submitted to detect differences in frequency of recalled idea units in different groups.
as main effect or (test versus game condition or high WMC versus low WMC condition) or interaction.

**Number of recalled idea units.** In general the number of free recalled idea units were similar across condition groups \(F(1,34)= 1.701, \text{ns; Eta}^2 = .048\), and across WMC \(F(1,34)= .957, \text{ns; Eta}^2 = .027\). However, the test condition partial Eta\(^2\) indicates a medium effect size. Readers in test condition recalled more \(M=15.52, \text{SD}=7.12\) than the readers in game condition \(M=13.00, \text{SD}=4.66\). The interaction effect was also not significant \(F(1,34)= .081, \text{ns; Eta}^2 = .002\). The mean scores and standard deviations across different factor groups for recalled idea frequencies presented separately in Table 4.

**The Effect of Reading Purpose and WMC on performance in the Application-Comprehension Measures**

It was hypothesized comprehension and application outcomes would differ as a function of WMC and reading purpose. Analysis of variance (ANOVA) was applied to detect differences performance of different groups on comprehension and application measures as main effect or (test versus game condition or high WMC versus low WMC condition) or interaction.

**Visual Comprehension Questions of the Game.** A two way ANOVA revealed that the performance on visual go questions did not differ across condition groups \(F(1,35)= 1.945, \text{ns; Eta}^2 = .053\), or WMC groups \(F(1,35)= .801, \text{ns; Eta}^2 = .022\). However, the partial Eta\(^2\) in test condition indicates a medium effect size. The readers in test condition performed better on visual game comprehension questions \(M=8.142, \text{SD}=1.74\) than the
readers in game condition (M=7.16, SD=2.33). The interaction effect was also not significant ($F(1,35)=.951, \text{ns; } \eta^2=.000$).

**Verbal Comprehension Questions of the Game.** A two way ANOVA revealed that the performance on verbal comprehension questions of the game did not differ across condition groups ($F(1,35)=.299, \text{ns; } \eta^2=.008$), or $WMC$ ($F(1,35)=1.126, \text{ns; } \eta^2=.031$). The interaction effect was also not significant ($F(1,35)=.001, \text{ns; } \eta^2=.000$).

**Application of the Go Game.** To determine which of the groups performed different on the application of the game, a two way ANOVA was run. The application scores did not differ across condition groups ($F(1,35)=2.827, \text{ns; } \eta^2=.075$), or $WMC$ groups ($F(1,35)=0.17, \text{ns; } \eta^2=.000$). However, the partial $\eta^2$ in test condition indicates a medium effect size. The readers in test condition performed better on the application of the game (M=7.33, SD=2.49) than the readers in game condition (M=6.16, SD=1.94). The interaction effect was also not significant ($F(1,35)=1.417, \text{ns; } \eta^2=.039$).

The mean scores and standard deviations across different factor groups for performance on the different comprehension application measures are presented separately in Table 5.

**Discussion**

The aim of this research was to explore how young readers adjust their comprehension processes as a function of reading purpose and working memory capacity. Elementary school readers involved in a reading and think aloud processes under different reading purpose conditions (Game versus Test). The main argument was that the conclusions from previous studies of the effect of reading purpose and WMC on reading comprehension have been limited for three reasons: (1) The assessment tools used in measuring WMC were not comprehensive and were not standardized; thus, it is was easy
to compare and find a meaningful pattern to make instructional decisions; (2) the existing research was extensive; however, the participants were adult readers, which may not be generalizable to young readers; (3) the ecological validity of the past research was disputable and needed further explanation.

Following the rationale of Linderholm and van den Broek (2002), we assessed young readers’ working memory capacity by the Working Memory Test Battery for Children (WMTB-C) subtests. This is a comprehensive measure of working memory designed for children. WMTB-C is devised to tap on all three components of working memory model; the phonological loop, visuo-spatial sketchpad and central executive proposed by Baddeley (2000). In addition to the standardized measures of working memory, vocabulary and reading comprehension measures were included. Finally, the participants were assessed on different reading comprehension questions reproduced from the text and the context and they played the game eventually.

The study demonstrated that reading purposes and working memory capacity did not produce any difference in this sample in utilization of different cognitive processes. The results were not statistically significant, with one exceptional case: The frequency of evaluative comments in game condition during free recall task significantly differed from the test condition. However, other differences in groups were examined in the light of the effect sizes.

The findings of this study supported our hypothesis that working memory capacity contributes to execute different cognitive processes in reading comprehension. The medium effect sizes in verbal protocol task revealed, the young readers with low WMC expressed more predictive and connecting-reinstatement inference than the readers
with high WMC. Linderholm and Van den Broek’s study (2002), found evidence that adult readers with low WMC uttered less predictive inference than readers with high WMC, and for separate categories of connecting and reinstatement inference frequency was similar for both WMC groups. Our findings are not in line with Linderholm and van den Broek’s (2002) findings. Moreover, in our study the low WMC readers demonstrated more predictive inference not only in verbal protocol task, but also in free recall task than high WMC readers. This finding was consistent within this study in both tasks, which might be an indication that low WMC readers are devoting more capacity to make predictions.

Van den Broek et al. (2001) also suggests that in order to maintain standards of coherence in comprehension, the readers are involved in inferential processes. It can be the case that High WMC group readers were more certain of their coherence building processes, therefore did not produce as much predictive and connecting-reinstatement inference as low WMC readers. On the other hand, not having the proper strategies caused our young readers’ to compensate their lack of understanding by producing inferences. The inconsistency of our results with Linderholm and van den Broek (2001) study can be explained by relation between age (college students versus elementary school students) and working memory capacity. Seigneuric and Ehrlich’s (2005) findings suggests that some processes in reading, for instance word recognition, is not automatized in early years of ages, and working memory capacity turns to be a determinant of reading comprehension after children become proficient in recognizing words. Therefore, the effect of age differences in both samples might cause the varying patterns in utterance of different cognitive processes.
The results also confirmed the prediction that different reading purpose conditions (game versus test) also caused readers to adjust their cognitive processes in different purpose conditions. The paraphrasing in test condition during the free recall task was more frequent than readers in game condition. This result is also in line with the previous findings of different investigations (Linderholm & van den Broek, 2002; Narvaez et al, 1999; van den Broek et al., 2001). Lorch et al. (1993) claim that school related tasks such as reading to study involves more re-reading and, thus, elicits more paraphrasing.

Apparently, both adult and young readers’ recall more in test condition, regardless of the different cognitive processes they generate or their working memory capacity. There is a tendency to recall more if the reading is for taking a test or studying for an exam. The test condition readers were also more effective in their recall than the game condition readers in this study. This evidence is perhaps pointing out that the frequency of the paraphrasing decides the performance on recall. Besides, the effect sizes in this research indicate that the recall was better for the young readers in test condition, as oppose to readers in game condition. This result shows that the current findings are in line with the two previous studies on the effect of reading purpose on text comprehension: Linderholm and van den Broek (2002) and van den Broek et al.(2001) reported that the results for free recall tasks were better among the adult readers in the study condition than in the entertainment condition. High level of comprehension is associated with paraphrases (Trabasso & Maliano, 1996). The sign of advanced grasp of the text might be due the effort to hold the unfamiliar information “in their own words”.

Although the cognitive processing patterns were similar across WMC and reading purpose groups, readers in test condition performed better on visual reading
comprehension questions of the game and in the application of the game than the readers in the game condition, as medium effect size indicated. This interesting finding illustrates that regardless of the readers’ working memory capacity, or the cognitive processes the test purpose condition causes young readers transfer their understanding of the concepts and rules to application context better than game condition readers. It might be the case that the young readers in test condition, perhaps because of the paraphrasing, recalled more and transferred more of their understanding into the comprehension and application tasks.

The results demonstrated that the reading purpose and working memory capacity was influential in the adjustment of young readers’ cognitive processes, and their recall for the text and application of what they read. The diversity in the findings of current study and the previous investigations in the literature might be for variety of reasons:

(1) van den Broek et al. (2005) suggest that comprehension cannot be quantified and measured easily, compare to other basic reading skills like vocabulary or phonological awareness. The incongruent results found in this study, might be due to the different assessment tools and methods used in various studies and in their scoring schemata.

The out loud protocols providing us a great deal of information about the comprehension processes, which otherwise would not be accessible to investigation. Although the consensus is that the think out loud protocols are the window on comprehension (Crain-Thoreson et al., 1997), there are still points to remark on the interpretation of the think out loud protocol data (Perfetti, 1985). The qualitative data provided by the out loud protocols might be undermined by different reasons involved in the speaking the ideas out loud. First of all, it is possible that participants do not speak
some of their idea units aloud. This is reflected on the frequency of the provided data. Next, as the participants are acknowledged that there will be no consequence for them, it might cause them feeling comfortable with generating incoherent or incomplete ideas. Furthermore, in the scoring and analysis part of the methodology the out loud protocols are very vulnerable to subjectivity of the examiners and might bias the results (Perfetti, 1985). Similarly, Crain-Thoreson et al. (1997) were concerned about out-loud protocols as a method of comprehension measurement:

“Perhaps the most intriguing result of our study was the demonstration that not all paraphrasing is equal. Although the most frequently used category in our original coding category, monitoring-comprehension, was significantly correlated with comprehension scores, less than 20% of the variance was accounted for by it. When we recorded the marked protocols and counted only those paraphrases and inferences that involved an active, accurate integration of passage content (transforming paraphrases), the percentage of the comprehension score variability accounted for rose to almost 49%.” (p. 588). Therefore, the reliability of out loud protocols is a critical issue in the interpretation of the results. For different conditions or particular age groups use of out loud protocols might maximally reflect the performance while in other case it may not be the best idea such as with old versus young readers, in different domains or different purposes etc. It can be speculated that the age difference in this study and the previous work might result in opposite findings. The adult readers are assumed to be equipped with tools to monitor their comprehension. If the readers are experienced in reading, it means that they have already developed their strategies to meet and sustain the criteria of coherence in
their understanding. Therefore, age might cause the differences in the findings in this present and the past findings.

Although one of the main aims of this study intended to obtain ecological validity, the accomplishment of this goal should to be elaborated: Another plausible explanation for non-significant results is context as a mediator. Readers, who were in grade 5 and 6 classrooms in the selected schools, were pulled out from their classrooms for participation in this study, for two sessions. The first session consisted of demanding standardized measures. The second session, which was at most one month apart from the first testing session, the experimenter applied vocabulary measure first and then the different reading purposes were introduced. The distinction was made clear by the experimenter in the testing sessions by verbalization of the purposes either; the participant was taking a test or playing a game after reading the text. Later, the participants were asked to apply what they have read in either game or test condition. In this context, we can speculate two different scenarios: First, the initial demanding standardized tests, the participants might have perceived the game condition also as a test, no matter what the examiner told them. If that was the case, then both groups would perform similarly in purpose conditions. The second scenario is that, the reading purpose had an effect on performance differences on comprehension and application measures for test condition group. It might be the case that the game condition cause the participants not to use their maximum capacity as it is only a game after all, why should they outperform? Weissberg and Paris (1986) investigated the effect of age and context (lesson versus game) on item recall among 3 to 7 years old children. Although the age group is not comparable with our sample, it might be given an idea of how students
conceptualize different conditions of lesson and game early on. The authors suggested that the game condition created uncertain task goals, distraction and processing load and this might suppress recall in young children. In the present study the readers in game condition produced different types of comprehension outcomes, namely more evaluative comments, connecting-reinstatement inferences and metacognitive comments. However, these cognitive processes they involved during reading did not aid them in transferring the information to the real life context when they needed to apply what they have just read. In sum, the testing context might have altered the results of this study, and caused the inconsistency in the findings of present and past research.

**Limitations**

There are several possible explanations for the failure to show a relationship between the effect of reading purpose and WMC on cognitive process adjustment in young readers. The issues related to the sample size, validity and reliability of the non-standardized measures, and procedures were reevaluated below:

1. Sample size was one of the most major limitations of this study. “How big the sample size should be?” has one simple answer with a common consensus in many resources: The larger the sample size, the better it is; because the larger sample sizes results in smaller sampling errors and more inclined to be generalized to the whole population (Johnson & Christensen, 2004; Murpy & Myors, 1998). The similar studies’ samples sizes range from 110 (Linderholm, & van den Broek, 2002) to 20 (Narvaez et al., 1999) participants. However, these studies involved college students either given credits for participation or they were called volunteers in undergraduate courses. This might have unemployed the genuineness of the participants’ performance. In short, the analyses
methods in this study were very conservative to find significant results with very small sample size. It may be likely that with a larger sample we might get significant results. In this study, we obtained non-significant results which lead us towards accepting that the different condition groups were indeed similar, and the results were due to chance. The time and resource constraints disabled the wish to involve more participants in this master’s thesis research project. Notwithstanding this, the data presented here can, however, be seen as preliminary evidence of a highly-specific analyses of how cognitive processes change in young readers with varying conditions of purpose and WMC. Even with the small sample size we had (N = 39), the ANOVAs detected medium to strong effect sizes for many outcome measures across reading purpose and working memory capacity groups. It was appropriate to discuss the results and the differences grounding on the effect sizes, otherwise, the likelihood of discarding results when they were true would increase.

(2) Reliability and validity of the measurement tools should be also reevaluated. Another reason to interpret the results of this study with caution is the use of one type of text for the tasks. The results obtained might be due to the text properties, rather than the effect of the independent variables. Even though, the difficulty level of the text was adjusted carefully, I may not be appropriate to differentiate the outcomes of comprehension. It might be also the case that the reading comprehension questions were not sufficiently demanding to distinguish the amount of comprehension in our sample.

(3) Perhaps the coding and scoring criteria were either too broad or more vulnerable to subjective judgments of the raters. The efforts in methodology to quantify the verbal responses might generate misleading results. First, the mutual understanding of the
categories by negotiation can minimize the controversial occurrences. However, if the category it is too broad, even though the inter-rater reliability is high enough for possible conclusions, it can be misleading amount of frequency in different categories. An interesting point on this matter was made by van den Broek et al. (2001). They explored how readers’ goals influence the inference generation and final recall. The researchers found out that readers in their study condition did not exhibit self monitoring as much as predicted. When they looked into their categories closely, they realized that the too broad categorization of comments led to misleading interpretations.

(4) The working memory capacity division method applied in this study was not the most desirable and accurate way to do so. In the original study, a reading span task was used to differentiate the WMC (Linderholm & van den Broek, 2002). They recorded the total number of words accurately recalled on this reading span task, and the upper and lower third of the frequency distribution of total words recalled by each participant was used to distinguish low- from high-WMC readers. Due to the small sample size in this study, discarding the data of participants could not be considered. Therefore, this division might come out as weakness in this present study.

Confounding effects of the first language was an important dimension, which should have been considered in initial in the set up. The participants were chosen from the multilingual city of Montreal, where the official languages are both English and French. In our sample, a big portion of our sample identified their first language both English and French. The sample received their education in English language, and their vocabulary level reflected a normal range for the sample in this study. However, the casual relationship between the first language and reading comprehension should be considered
initially in a similar study in the future. In a study, participants whose first language were not English and read very similarly and did not adjust processing across reading purpose conditions. Readers with the lowest English proficiency may not have kept a reading purpose in mind as reading proceeded because it put too much strain on working memory resources. In the case when the reading task is demanding, readers with limited resources may not be able to maintain the reading purpose (Horiba, 2000), because it put too much taxing for working memory capacity. Similarly, Linderholm and van den Broek (2002) argued that maintaining reading purpose during reading may be more difficult when the task is more complex or demanding for the reader who has less capacity to process and hold information. Thus, replicating the current study with a more homogeneous language background might alter our findings.

Next, the reader’s standard of coherence is highly related to their reading strategies. On one hand, standards of the individual give the reader a feeling that comprehension is achieved. On the other hand the readers’ metacognitive skills determine if the coherence can be attained. The metacomprehension skills are directly involved in how readers adjust to diverse reading requirements (Lorch, Lorch, & Klusewitz, 1993). Although in this research measuring metacognitive skills was not a focus, but it might have had an effect on the production of cognitive processes. We controlled the effect of prior knowledge by exploring if the readers were familiar with the game presented in the tasks. However in the future, role of metacognitive strategies and skills in young readers could advance our understanding of how comprehension processes are adjusted by employment of different strategies in varying reading purposes conditions. The outcome of a successful reading is a coherent mental representation of the text. The readers, especially
the young ones may not realize that the coherence is achieved due to their lack of metacognitive strategies (van den Broek, 2004). If the reader is experienced in reading, they might have already developed strategies monitor coherence. Additionally, highly skilled readers have strategic tools in their repertoire, which enable them to learn from the texts even they have low levels of prior knowledge (Crain-Thoreson et al., 1997). Since in our sample, the readers were as young as 11 years old, we cannot assume that they have basal level of reading skills as adults. Therefore, age might cause the differences in the findings in this present research and the past findings. Obviously, the quantity of the processes does not tell us the whole story about comprehension. As Narvaez et al. (1999) argued the response generation might be qualitative rather than quantitative. Meeting the requirements of coherent understanding of the text, such as making correct inferences, accurately connecting information or paraphrasing or creating true representation, is desirable in real life context rather than only manufacturing idea units. Therefore, I suggest that future studies should investigate how the qualities of cognitive processes differ in young readers. In sum, the current research study demonstrated that young readers adjust their cognitive processes according to reading purpose and working memory capacity. In addition, the reading purpose affects the final free recall in different conditions. These findings are in line with the existing literature, and contribute to show evidence in how reading purpose and capacity to hold and manipulate information simultaneously triggered and differentiated comprehension and application skills in young readers. Exploring the accuracy of recall and comprehension processes in a similar design, but with a larger sample size might be another research agenda.
References


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


Table 1

**Means, Standard Deviations for the WMTB-C Subtests, Global Working Memory, PPVT-III and WJ-Passage Comprehension by Reading Purpose Group**

<table>
<thead>
<tr>
<th>Tests</th>
<th>Test M</th>
<th>SD</th>
<th>N</th>
<th>Game M</th>
<th>SD</th>
<th>N</th>
<th>Total M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT-III</td>
<td>102.05</td>
<td>13.88</td>
<td>21</td>
<td>97.11</td>
<td>9.47</td>
<td>18</td>
<td>99.77</td>
<td>12.15</td>
<td>39</td>
</tr>
<tr>
<td>WJ Passage Comp.</td>
<td>90.86</td>
<td>6.68</td>
<td>21</td>
<td>88.11</td>
<td>6.02</td>
<td>18</td>
<td>89.59</td>
<td>6.45</td>
<td>39</td>
</tr>
<tr>
<td>Word List Recall</td>
<td>123.43</td>
<td>15.62</td>
<td>21</td>
<td>114.39</td>
<td>18.76</td>
<td>18</td>
<td>119.26</td>
<td>17.51</td>
<td>39</td>
</tr>
<tr>
<td>Nonword List Recall</td>
<td>126.38</td>
<td>18.64</td>
<td>21</td>
<td>125.06</td>
<td>15.98</td>
<td>18</td>
<td>125.77</td>
<td>17.25</td>
<td>39</td>
</tr>
<tr>
<td>Block Recall</td>
<td>115.48</td>
<td>15.15</td>
<td>21</td>
<td>118.78</td>
<td>19.75</td>
<td>18</td>
<td>117.00</td>
<td>17.27</td>
<td>39</td>
</tr>
<tr>
<td>Listening Recall</td>
<td>127.24</td>
<td>11.30</td>
<td>21</td>
<td>128.67</td>
<td>11.13</td>
<td>18</td>
<td>127.90</td>
<td>11.09</td>
<td>39</td>
</tr>
<tr>
<td>Counting Recall</td>
<td>115.67</td>
<td>8.30</td>
<td>21</td>
<td>113.33</td>
<td>14.25</td>
<td>18</td>
<td>114.59</td>
<td>11.34</td>
<td>39</td>
</tr>
<tr>
<td>Global Working Memory</td>
<td>121.64</td>
<td>8.68</td>
<td>21</td>
<td>120.04</td>
<td>9.10</td>
<td>18</td>
<td>120.90</td>
<td>8.79</td>
<td>39</td>
</tr>
</tbody>
</table>

Note. Except PPVT-III, all measures are subtests of the measures

Table 2

**Means and Standard Deviations of Utterance Frequency of Different Cognitive Processes during Verbal Protocol Task as a Function of Reading Purpose and WMC.**

<table>
<thead>
<tr>
<th>Cognitive Processes Response Categories</th>
<th>Test M</th>
<th>SD</th>
<th>p</th>
<th>Game M</th>
<th>SD</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associations</td>
<td>.37</td>
<td>.55</td>
<td>ns</td>
<td>.44</td>
<td>.72</td>
<td>ns</td>
</tr>
<tr>
<td>Evaluative Comments</td>
<td>.75</td>
<td>1.22</td>
<td>ns</td>
<td>.67</td>
<td>1.32</td>
<td>ns</td>
</tr>
<tr>
<td>Cognitive Processes Response Categories</td>
<td>Test</td>
<td>Game</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>------------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>p</td>
<td>M</td>
<td>SD</td>
<td>p</td>
</tr>
<tr>
<td>Associations</td>
<td>.22</td>
<td>.44</td>
<td>ns</td>
<td>.25</td>
<td>.46</td>
<td>ns</td>
</tr>
<tr>
<td>Evaluative Comments</td>
<td>.89</td>
<td>1.36</td>
<td>ns</td>
<td>1.50</td>
<td>1.77</td>
<td>ns</td>
</tr>
<tr>
<td>Elaborative Inferences</td>
<td>1.11</td>
<td>1.05</td>
<td>ns</td>
<td>1.38</td>
<td>1.30</td>
<td>ns</td>
</tr>
<tr>
<td>Predictive Inferences</td>
<td>3.33</td>
<td>3.35</td>
<td>ns</td>
<td>4.25</td>
<td>2.49</td>
<td>ns</td>
</tr>
<tr>
<td>Connecting-Reinstatement Inferences</td>
<td>2.33</td>
<td>2.87</td>
<td>ns</td>
<td>4.75</td>
<td>3.77</td>
<td>ns</td>
</tr>
<tr>
<td>Metacognitive Comments</td>
<td>3.11</td>
<td>2.09</td>
<td>ns</td>
<td>2.50</td>
<td>2.56</td>
<td>ns</td>
</tr>
<tr>
<td>Paraphrases</td>
<td>19.67</td>
<td>5.24</td>
<td>ns</td>
<td>16.38</td>
<td>5.13</td>
<td>ns</td>
</tr>
</tbody>
</table>

*Note.* Numbers indicate average frequency of idea units expressed, per participants in each condition. WMC=Working Memory Capacity, p > .05, ns = non-significant

Table 3

*Means and Standard Deviations of Cognitive Processes Performed during Free Recall Task as a Function of Reading Purpose and WMC.*
<table>
<thead>
<tr>
<th>Cognitive Processes</th>
<th>Test</th>
<th>Game</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>M</em></td>
<td><em>SD</em></td>
</tr>
<tr>
<td>Associations</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluative Comments</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>Elaborative Inferences</td>
<td>.75</td>
<td>.75</td>
</tr>
<tr>
<td>Predictive Inferences</td>
<td>.77</td>
<td>.73</td>
</tr>
<tr>
<td>Connecting-Reinstatement Inferences</td>
<td>1.42</td>
<td>1.24</td>
</tr>
<tr>
<td>Metacognitive Comments</td>
<td>.50</td>
<td>1.17</td>
</tr>
<tr>
<td>Paraphrases</td>
<td>10.25</td>
<td>5.26</td>
</tr>
</tbody>
</table>

**Low WMC Group**

<table>
<thead>
<tr>
<th>Cognitive Processes</th>
<th>Test</th>
<th>Game</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>M</em></td>
<td><em>SD</em></td>
</tr>
<tr>
<td>Associations</td>
<td>.11</td>
<td>.33</td>
</tr>
<tr>
<td>Evaluative Comments</td>
<td>.11</td>
<td>.33</td>
</tr>
<tr>
<td>Elaborative Inferences</td>
<td>.44</td>
<td>.53</td>
</tr>
<tr>
<td>Predictive Inferences</td>
<td>1.13</td>
<td>1.29</td>
</tr>
<tr>
<td>Connecting-Reinstatement Inferences</td>
<td>1.78</td>
<td>1.86</td>
</tr>
<tr>
<td>Metacognitive Comments</td>
<td>1.56</td>
<td>1.33</td>
</tr>
<tr>
<td>Paraphrases</td>
<td>10.22</td>
<td>4.55</td>
</tr>
</tbody>
</table>

*Note.* Numbers indicate average frequency of idea units expressed, per participants in each condition. WMC=Working Memory Capacity, p* < .04
### Table 4

**Means and Standard Deviations of Free Recalled idea units as a Function of Reading Purpose and WMC.**

<table>
<thead>
<tr>
<th>Group</th>
<th>Test Game Total</th>
<th>Test Game Total</th>
<th>Test Game Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>p</td>
</tr>
<tr>
<td>High WMC Readers</td>
<td>14.42</td>
<td>7.25</td>
<td>ns</td>
</tr>
<tr>
<td>Low WMC Readers</td>
<td>17.00</td>
<td>7.09</td>
<td>ns</td>
</tr>
<tr>
<td>Total</td>
<td>15.52</td>
<td>7.52</td>
<td>ns</td>
</tr>
</tbody>
</table>

*Note.* Unit of measurement is the mean number of idea units expressed, per reader in each condition. WMC=Working Memory Capacity, p > .05, ns = non-significant

### Table 5

**Means and Standard Deviations of scores on the different comprehension application measures as a Function of Reading Purpose and WMC.**

<table>
<thead>
<tr>
<th>Comprehension-Application Measures</th>
<th>Test</th>
<th>Game</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Visual-Reading Comprehension Questions</td>
<td>8.41</td>
<td>1.62</td>
</tr>
<tr>
<td>Verbal Reading-Comprehension Questions</td>
<td>7.91</td>
<td>1.44</td>
</tr>
<tr>
<td>Application of the Go Game</td>
<td>6.91</td>
<td>2.02</td>
</tr>
<tr>
<td>Low WMC Group</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comprehension-Application Measures</th>
<th>Test</th>
<th>Game</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Visual-Reading Comprehension Questions</td>
<td>7.77</td>
<td>1.92</td>
</tr>
<tr>
<td>Task</td>
<td>Mean 1</td>
<td>Mean 2</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Verbal Reading-Comprehension Questions</td>
<td>7.44</td>
<td>.88</td>
</tr>
<tr>
<td>Application of the Go Game</td>
<td>7.88</td>
<td>3.05</td>
</tr>
</tbody>
</table>

*Note.* Numbers indicate average frequency of idea units expressed, per participants in each condition. WMC=Working Memory Capacity, p < .05
Appendices

Appendix A: *The text*

How to Play the Ancient Game “Go” by Karl Baker

Go is a board game that was first played in China 4000 years ago. People play this game in the Far East, and it is becoming more and more popular in Europe and in the United States. It is the oldest game in the world and is still played by millions of people. “Go” is one of the Japanese names for the game; it means “the surrounding game”. It is a game of skill and players must use strategies to win.

What are the goals of the game?

There are four goals in the game: (1) Surround territory (2) Reduce your partner’s territory (3) Capture your partner’s stones (4) Protect your own stones.

How does Go work?

Two players compete with each other to gain territory by placing stones on a smooth wooden board with a simple grid drawn on it, usually 19 by 19 lines. Play begins on an empty board. Each player wants to surround territory with his or her stones. Each player wants to control and capture more territory than the other. The purpose of the game is to capture the most territory.

The board has horizontal and vertical lines that crisscross, much like a chessboard except that you play on the intersections of the lines instead of on the squares. You may also play on the outer edges and corners of the board.

On the board, where the lines intersect is called a point. Each point is a valuable piece of territory in this game. The object of the game is to completely surround or fence in as many points of territory as you can.
The markers of play are called stones, of which one set is black and the other is white. One player takes white while the other takes black. Black goes first, and the players take turns playing one stone each on one of the empty points. Once a stone is placed on the board, it does not move to another point.

### Forming Connections

Each point on the board has another point beside it, which is called a nearby point. When another stone of the same color is placed on a nearby point, the two stones are called connected unit. Once connected, all stones of same color form an inseparable unit. Units can be made of one or more stones. Stones of the same color that sit beside each other along a line are treated as a single unit.

Each nearby point of a unit is called a liberty, if it is empty. Remember, the lines go horizontally and vertically, and a point which is diagonally beside a point is not its nearby point, because it is not connected by one straight line.

### Capturing Your Partner’s Stones

The player should try to place their stones on the liberties of the other player's stones. In other words, to prevent your partner from taking more territory than you, you should try to place your own stones at liberties which are immediately beside his or her stones. If you are successful in surrounding your partner’s stones you will have captured them. Captured stones are immediately removed from the board and are kept as prisoners. Players spend much of the time trying to find escape way for their own stones and to prevent the escape of their partner’s stones.

If a unit of stones does not link up to any empty points, the stone or group is considered captured and are removed from the board.

As you begin to play Go, you should warn your partner when you are able to capture their stones by saying the word “atari.” Saying “atari” means: “I can capture your stones on my next turn.”

### Some Tactics for Effective Play

1. Here are some guidelines for playing Go more successfully:
2. Keep your stones connected to each other; it makes them more difficult to be captured.

3. Avoid making lots of groups; do not scatter your stones around.

4. Look for weak groups in your partner’s territory, that is, groups that are connected to relatively few empty intersections. Attack the other player’s weakest group. Make your own weak groups stronger by attaching more stones to them.

5. Try to surround empty intersections with walls (make “forts”) so as to have more safe places to play.

6. Play outside your forts as long as you can safely do so, saving your inside places for later.
Appendix B: Visual Comprehension Questions

Practice Question 1: How many liberties does this stone have?

Practice Question 2: There are three places you can choose to put your black stone on either A, B, or C. At which point must black place a stone in order to capture the white stone?

Question 1: How many points are there in the example below?

Question 2: How many liberties does this black stone have?
Question 3: Again, I put some letters at certain points. At which point must black place a stone in order to capture the white stone and remove the unit from the board: A, B, C or D?

Question 4: At which point must black place a stone in order to capture the white stone and remove it from the board?

Question 5: At which point must black place a stone in order to capture the white stone and remove it from the board?

Question 6: At which point must black place a stone in order to capture the white stones and remove it from the board?
Question 7: At which point must black place a stone in order to eventually capture the white stone and remove it from the board?

Question 8: The following figure shows the position after the black player took his or her turn. Where do you think the white stones were located before they were captured and removed from the board?

Question 9: Suppose that you are black, at which point can black play in order to rescue the five-stone black unit: A, B, C or D?
Question 10: How many *liberties* does this unit have in total?

![Diagram of Go game board]

Appendix H: Verbal Comprehension Questions of Game Go

1. What does “Go” mean in Japanese?
   a) take and never give back game
   b) surrounding game
   c) defeat game

2. In which part of the world is the game of Go played?
   a) Far east, America, Europe
   b) France, Japan, China
   c) Both a and b

3. Which one below is not a goal in the game of Go?
   a) Try to get the black stones when you start
   b) Capture your partner’s stones
   c) Protect your own stones

4. Imagine that you are playing this game with a friend, how do you decide who starts first?
   a) Rock, Paper & Scissors
   b) The color of the stones decides who starts first
   c) I give my partner the chance to start first

5. What could you do to make your stones stronger?
   a) Scatter your stones around the board.
   b) Keep your stones connected to each other
   c) Avoid getting too close to your partner’s stones.

6. When you say __________ to your partner it means: “I can capture your stones on my next turn.”
7. True / False Once a stone is placed on a point, you have to move your stones to different points in order to make connections.

8. True / False You can make your own Go game materials.

9. True / False the objective of the game is to completely surround or fence in more points of territory than the person you are playing with.

10. True / False Captured stones are immediately removed from the board and are kept as prisoners.

Appendix D: Ethics Approval Form
Faculty of Education – Review Ethics Board
Certificate of Ethical Acceptability of Research Involving Humans

REB File #: 754-1206

Project Title: Influence of working memory capacity and reading purpose on children’s reading comprehension

Applicant’s Name: Z. Ozlem Cankaya

Department: ECP

Status: Master’s student Supervisor’s Name: Robert Savage

Granting Agency and Title (if applicable): n/a

Type of Review: Expedited ✓ Full

This project was reviewed by: Starke-Meyerring/Shariff

Approved by

[Signature]
Robert Bracewell, Ph.D.
Chair, Education Ethics Review Board

Approval Period: Dec. 11, 2006 to Dec. 11, 2007

All research involving human subjects requires review on an annual basis. An Annual Report/Request for Renewal form should be submitted at least one month before the above expiry date. If a project has been completed or terminated for any reason before the expiry date, a Final Report form must be submitted. Should any modification or other unanticipated development occur before the next required review, the REB must be informed and any modification can’t be initiated until approval is received. This project was reviewed and approved in accordance with the requirements of the McGill University Policy on the Ethical Conduct of Research Involving Human Subjects and with the Tri-Council Policy Statement on the Ethical Conduct for Research Involving Human Subjects.

12/11/06