A LATENT VARIABLE ANALYSIS OF WORKING MEMORY AND READING COMPREHENSION: EXAMINING THE ROLE OF ATTENTION CONTROL AND LONG-TERM MEMORY RETRIEVAL IN COMPREHENSION PERFORMANCE

by

GREGORY JOSEPH SPILLERS

(Under the Direction of Paula Schwanenflugel)

ABSTRACT

Working memory consistently predicts reading comprehension; however, the mechanisms underlying this relationship are still poorly understood. While many researchers have argued that short-term attentional processing is the defining factor, there is evidence to suggest that the ability to retrieve information from long-term memory may be just as important. The current study investigated the relationship between working memory and reading comprehension and the specific control processes that underlie this relationship using latent variable analysis. Multiple measures of working memory capacity, attention control, long-term memory retrieval, and reading comprehension were employed to create latent factors of each construct. Confirmatory factor analysis and structural equation modeling revealed that it was long-term memory retrieval, and not attention control, that predicted variance in reading comprehension performance. Furthermore, a sub-group analysis confirmed that reading comprehension is almost exclusively dependent on the ability to retrieve information successfully from long-term memory, with little contribution from attention control abilities.

INDEX WORDS: WORKING MEMORY, LONG TERM MEMORY, ATTENTION CONTROL, READING COMPREHENSION

A LATENT VARIABLE ANALYSIS OF WORKING MEMORY AND READING COMPREHENSION: EXAMINING THE ROLE OF ATTENTION CONTROL AND LONG-TERM MEMORY RETRIEVAL IN COMPREHENSION PERFORMANCE

by

GREGORY JOSEPH SPILLERS

B.A., University of North Carolina, Greensboro, 2008

M.S., University of Georgia, 2010

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2017

A LATENT VARIABLE ANALYSIS OF WORKING MEMORY AND READING COMPREHENSION: EXAMINING THE ROLE OF ATTENTION CONTROL AND LONG-TERM MEMORY RETRIEVAL IN COMPREHENSION PERFORMANCE

by

GREGORY JOSEPH SPILLERS

Major Professor: Paula Schwanenflugel

Committee: Martha Carr
Amanda Ferster
Logan Fiorella

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2017

DEDICATION

To Miranda, to whom I owe everything.

ACKNOWLEDGEMENTS

I want to thank Dr. Paula Schwanenflugel, the members of my committee, and the Department of Educational Psychology for providing me with the resources to conduct the research presented in the following pages and the encouragement to tackle the problems in which I was most interested. I would also like to acknowledge the support of Dr. Nash Unsworth, to whom I am very fortunate to call both a mentor and friend. Finally, I wish to thank the University of Georgia, both specifically for the Tuition Assistance Program, and generally, for taking the chance on a young Tarheel nearly a decade ago and granting him every opportunity to find success.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER 1

BACKGROUND

For the last several decades, working memory has been a useful concept in psychology for explaining how individuals simultaneously (and successfully) process, store, and manipulate complex information and ideas. Beyond its explanatory power, the importance of working memory as a construct in cognitive psychology also lies in its relation to other higher-order cognitive abilities. Over the years, volumes of research has been conducted showing that measures of working memory strongly predict performance on a host of other tasks, such as measures of intelligence (Ackerman, Beier, & Boyle, 2005; Engle & Kane, 2003; Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle, 2004; Unsworth, Brewer, & Spillers, 2009), vocabulary learning (Daneman & Greene, 1986), mathematical ability (Swanson & Beebe-Frankenberger, 2004; Geary, Hoard, Byrd-Craven, & Desoto, 2004; Bull & Scerif, 2001), and performance on standardized tests such as the SAT and ACT (Jurden, 1995; Engle, Tuholski, Laughlin, & Conway, 1999; Turner & Engle, 1989). Further, working memory has also been found to be an important predictor of other behavioral phenomena such as susceptibility to choking under pressure (Beilock & Carr, 2005), particular personality traits (Unsworth, Miller, Lakey, Young, Meeks & Campbell, 2009), and even early onset Alzheimer’s (Rosen, Bergeson, Putnam, Harwell, & Sunderland, 2002).

Given the utility of working memory in explaining how individuals encode information and execute a variety of complex cognitive operations, it should come as no surprise that it has also been extensively employed in discussions of the learning process in both children and adults, specifically as it pertains to the development of reading and comprehension skills (Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Daneman & Carpenter, 1980; 1983; Daneman & Merikle, 1996; Moulin & Gathercole, 2007; Nevo & Breznitz, 2011; 2013; Swanson, Howard, & Sáez, 2006). Reading comprehension is a pre-requisite for both academic achievement and success in later life. Students who fall behind their peers in reading are often unable to catch up and this has been found to have a deleterious, long-term effect on future academic achievement (Alloway, Gathercole, Kirkwood, & Elliott, 2009; Dahlin, 2011). Moreover, the act of reading and understanding what has been read is a complex mental process that requires the storage and coordination of many elements, including letter-sound correspondences, word recognition, syntactical and semantic information, etc. Readers must manage all of these features and coordinate them seamlessly for effective comprehension. It would make sense then, that working memory, the primary system argued to handle effortful mental operations, has been found to be one of the strongest predictors of reading comprehension performance. In fact, researchers have found that low working memory skills constitute a high risk factor for educational underachievement across the primary school years and that difficulties in reading comprehension can be directly linked to deficits in working memory capacity (Cain, Oakhill, Bryant, 2004; Swanson, 1999; Sesma, Mahone, Levine, Eason, & Cutting, 2009**;** Locascio, Mahone, Eason, & Cutting, 2010). Importantly, researchers have found that even after other general abilities (such as vocabulary knowledge, IQ, phonological awareness, and verbal short-term memory) have been statistically accounted for, working memory skills still significantly predicted performance on particular reading measures, suggesting that the relationship between working memory and reading difficulties is independent of related cognitive skills (Sesma et al., 2009; Swanson, 2003; Swanson & Saez, 2003; Alloway et al., 2009).

Though close links between working memory function and reading comprehension ability have been consistently shown, the specific mechanism(s) underlying this relationship has depended in large part on the conception of working memory to which a given researcher or practitioner subscribes. Indeed, the nature and purpose of working memory has been a key problem in cognitive psychology since the early days of the field, and our understanding of the construct has evolved over the decades from simple box models inspired by “the brain as a computer” metaphors to distributed neural networks provoked by the advent of neuroimaging technology (see Miyake & Shah, 1999; Spillers, Brewer, & Unsworth, 2012). Despite these advances in our understanding of working memory, reading comprehension researchers have been reluctant to incorporate newer ideas about the construct into their explanations for why working memory is such a strong predictor of reading comprehension performance. In fact, even the most recent work examining the role that working memory plays in reading comprehension incorporates ideas and models several decades old. Certainly, this firm reliance on certain models is a testament to the explanatory power of those models but it may also be detrimental to our goal of constructing a comprehensive understanding of the relationship between working memory and reading comprehension.

The next few sections will summarize developments in research on the nature of working memory and discuss the impact these developments have had on our understanding of working memory’s relationship to reading comprehension. Following this review, a new study will be proposed aimed at investigating the underlying mechanisms responsible for the relation between working memory to reading comprehension using latent variable analysis.

The beginnings of working memory: The modal model

The initial concept of working memory was borne out of growing dissatisfaction with the earlier concept of short-term memory; that is, a system used for the sole purpose of storing some amount of information for a limited time. Researchers felt that a short-term store was far too rigid and that such a passive store was insufficient for explaining how individuals successfully execute the many complex activities required of them on a daily basis. Thus, there was a need for a new type of construct that was flexible and dealt with both the storage and manipulation of information in the interim. Although Miller, Galanter, and Pribram (1960) are often credited for coining the term working memory, it was Atkinson and Shiffrin (1968)'s treatment of the construct that first laid foundation for the modern conception. In the Atkinson-Shiffrin model of memory (or modal model), information passes linearly through three separate stages (or stores): a sensory store, a short-term store, and a long-term store (see Figure 1). This "pipeline" model of memory was not new (see Broadbent, 1958), however, Atkinson and Shiffrin recognized the empirical pitfalls accompanying a fixed short-term store, so they paired with it multiple control processes, dubbing this new dynamic construct a person's working memory—a system, they argued "in which decisions are made, problems are solved and information flow is directed" (Atkinson & Shiffrin, 1971, p. 83). These control processes included rehearsal, coding, and retrieval and represented the first detailed attempt to distinguish a *flexible control system* from static memory stores. Rehearsal was argued to be the primary process through which information is both maintained in the transient short-term store and transferred to the more permanent long-term store. In the absence of rehearsal, forgetting occurs as traces are argued to decay, usually within 30 s. Coding referred to the type of information that is transferred (i.e., what aspects of the information being processed are registered in the long-term store). Retrieval was important for getting the information already stored in the long-term store back into working memory for current processing or output.

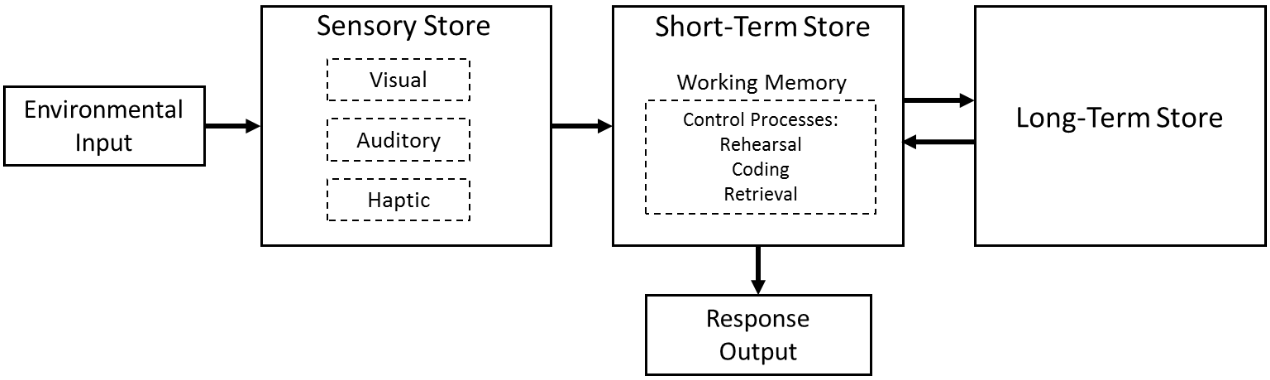


Figure 1. Schematic representation of the modal model of memory. Adapted from Atkinson & Shiffrin (1971).

Reading comprehension and the modal model of memory

Though comparatively simple to how we conceive of working memory today, Atkinson & Shriffin's modal model was robust enough to provide researchers with a mechanism for describing how individuals learn and process information. In fact, work by Kintsch and colleagues (Kintsch, 1974, 1988, 2004; Kintsch & Vipond, 1978; Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983) was some of the first to utilize this model of memory for explaining how discourse and text information is processed. While their primary aim was to explore how one might compute the readability of text and the factors that affect how easily and how much an individual comprehends what they are reading, the authors described a process that incorporates the idea that information flow is first directed through a limited short-term store, where it is acted upon by certain control processes, and then transferred into long-term memory.

Briefly, Kintsch and colleagues developed a framework for constructing meaning from written text entitled the Construction-Integration (C-I) theory of reading comprehension. In this theory, the act of comprehension is viewed as occurring on multiple levels. Specifically, readers develop a representation of a text from the bottom-up—first, constructing and then integrating lower-level text information (syntax, semantics, etc.) with their prior knowledge. The initial level involves developing a strictly perceptual representation of the text in terms of the surface form or exact words and phrasing that appear, called the *surface code*. At the second level, the surface code is updated to form the *text-base*, a slightly more processed level of representation in which the information that was read is now represented as a basic set of overlapping propositions that interconnect ideas derived from the text. These propositions are simply basic units of meaning abstracted from the words that have been read. Importantly, the text-base representation includes not only information derived strictly from what is being read but also incorporates associations from long-term memory (i.e., prior knowledge) and minor inferences drawn from the text content. As the reader progresses through the text, these propositions and associations are updated, strengthened, and elaborated upon. This process leads to integration of new information with the reader's existing knowledge base; thereby, forming a coherent, abstract representation of the text, called the *situation model*. This level of representation is now no longer explicitly text-based but becomes the instantiation of both explicit and implicit relations between ideas, such that a reader should now be able to flexibly describe what was read, establish causal and logical connections, and make inferences for use in novel contexts. Kintsch and colleagues argue that creation of the situation model can be considered as *the* process of deeply understanding a text, or comprehension. As the number of propositions increase in a given text, so does the text's overall complexity.

With the parameters of the Construction-Integration model firmly established, Kintsch and colleagues then employed the modal model as the mechanism through which the construction and integration process occurs. Specifically, they proposed that individuals begin to store propositions in a temporary buffer (akin to working memory) as they read through the text. These propositions are argued to form a chunk that is processed, retained, and connected with new incoming information as the reader moves through the text. Chunks are argued to be necessary given the limited constraints of this temporary buffer and the "size" and integrity of a given chunk depends on both text and reader characteristics (predicting, as we will see, much of the latter work involving individual differences in working memory capacity). As the buffer is limited, some propositions are kept in a processing chunk, while others fall out. Overlap between the chunked propositions and incoming material assists in creating understanding and ensuring comprehension (or, as Kintsch and colleagues describe it, "coherence"). Importantly, if a connection cannot be made between the new input and the retained propositions, a search of previously processed propositions from long-term memory is conducted to determine whether there is an already stored proposition that could be connected with the information retained in working memory (hence, why novel propositions are considered more difficult to comprehend compared with those that have been repeated). If not, the reader must create an inference to connect the information.

The above model denotes several facets about the relationship between working memory and reading comprehension. First, the more often a set of propositions is input into working memory (or buffer), the greater the chance it will be stored in long-term memory. This implies that working memory is a necessary conduit for memory representations to be stored intact. Second, the above model also suggests that comprehension occurs in the short-term (or working memory) buffer. Specifically, propositions are temporarily stored in chunks while previously stored propositions (derived either from the current text or prior knowledge) are reinstated from long-term memory. Finally, and on a related note, the model implies comprehension is the result of a parallel process between both working memory and long-term memory. Text is read, stored temporarily while being processed, and then compared to previously processed information that has been simultaneously reinstated (or reactivated). Information is being shuffled into and out of working memory and long-term memory concurrently. This idea that short-term (or working) memory operates in conjunction with long-term memory, although interesting, is one that, as will be seen, becomes less emphasized in favor of understanding the limits or capacity of working memory and how these limitations dictate performance on certain cognitive activities, including comprehension.

The multi-component model of working memory

The work of Atkinson and Shiffrin (1968) represented an obvious advancement in how memory researchers conceptualized short-term memory (and the human memory system more broadly), however, it was not long before researchers began proposing alternatives. Arguably the most celebrated of these alternatives is the multicomponent model of working memory first proposed by Baddeley and Hitch (1974; see also Baddeley, 1986). In their seminal chapter in the Psychology of Learning and Motivation, Alan Baddeley and Graham Hitch proposed a model of working memory characterized by multiple stores, each with a capacity for processing specific kinds of information. In its original formulation, the Baddeley and Hitch model of working memory contained three distinct components: two slave systems, the *phonological loop* and the *visuo-spatial sketchpad*, both of which were headed by a supervisory, master system, the *central executive*. As suggested by their names, the phonological loop and visuospatial sketchpad were responsible for the storing and processing of auditory and visual content, respectively. The central executive served as a coordinator of the subsidiary components, organizing the interplay between the two buffers and long-term memory. Its primary functions included the allocating of attention to the slave systems when necessary, planning future actions and behaviors to solve problems, and the control of information processing, more generally. More recent work by Baddeley (2000) has seen the addition of a fourth component, the *episodic buffer*, argued to provide temporary storage of information held in a multimodal code, which is capable of binding information from the subsidiary systems and long-term memory into a unitary episodic representation. The basis for the multi-component model arose out of work examining the costs associated with doing two tasks concurrently—the logic being that if two tasks rely on the same processing mechanisms, they should show interfering effects on one another. Baddeley and Hitch found that two tasks interfered with each other only to the extent that they were of similar kind. Specifically, they showed a double dissociation: Verbal tasks interfered with verbal short-term memory but not visual short-term memory, and visual tasks interfered with visual short-term memory but not verbal short-term memory, thus, giving reason for separate buffers responsible for different modes of information. Baddeley and Hitch's working memory model has been both an empirical and theoretical tour de force and is still widely considered the single most influential model of working memory among many researchers in the fields of cognitive psychology and neuroscience today.

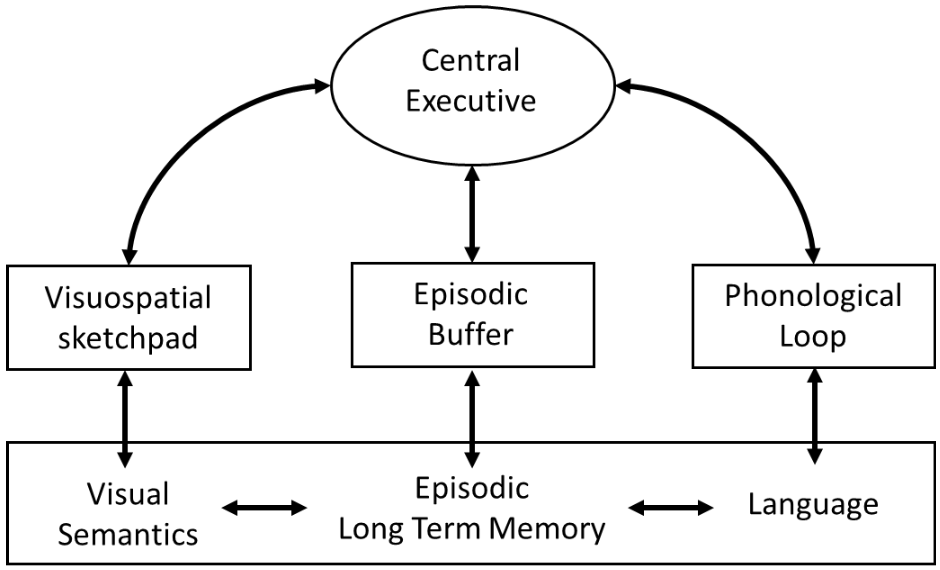


Figure 2. Schematic representation of the multicomponent model of working memory. Adapted from Baddeley (2000).

Reading comprehension and the multi-component model of working memory

Shortly after proposing the multi-component model, Baddeley and colleagues set to work to test the model in a number of different domains. Not surprisingly, initial efforts to use the model to explain reading and comprehension focused on the role of the phonological loop. As mentioned above, the phonological loop is responsible for receiving, temporarily storing and processing acoustic, speech, and some visual-based information. Further, the phonological loop system has been delineated into two-subcomponents: the short-term phonological store (the component that actually stores information) and the articulatory loop (a sub-vocalization rehearsal process argued to keep information active in the phonological store). Notably, both the system, generally, and its subcomponents have been heavily implicated in the acquisition and development of language (Baddeley, 1979; 1992; Gathercole & Baddeley, 1993a; 1993b).

In an attempt to define the specific aspects of the phonological loop that affect comprehension performance, Baddeley and Hitch (1974) conducted a series of experiments examining the influence of the phonological store and the articulatory loop, respectively. In one such study, they directly tested the effect of concurrent memory load in the phonological store on comprehension by having participants simultaneously store a set of random digits (with variable set-size conditions of 3-6) while listening to an experimenter read short passages derived from the commonly employed Neale Analysis of Reading Ability. Following passage presentation, the participants were allowed to abandon the digit task and answered several comprehension questions. Consistent with earlier findings in list-learning paradigms, comprehension (as measured by the proportion of questions answered correctly) was significantly worse as digit load increased. In another study designed to examine the role of the articulatory loop, Baddeley and Hitch (1974) tested the influence phonemic similarity would have on an individual's ability to process and comprehend text. In this study, participants were asked to read a series of short sentences and judge whether each sentence was sensible or nonsense. Further, in some trials, the sentences were made of phonemically similar words ("Red headed Ned said Ted fed in bed) while in other trials words were phonemically dissimilar ("Dark skinned Ian thought Harry ate in bed"). Sensible vs. nonsense trials were created by switching the order of two words in each sentence ("Red headed Ned Ted said fed in bed"). Baddeley and Hitch found that the mean decision times for phonemically similar sentences (both sensible and nonsense) were higher compared with phonemically dissimilar. Thus, they concluded that not only was the amount of information a reader could reliably maintain in the phonological store important for adequate reading comprehension but that the ability for an individual to use sub-vocalization (via the articulatory loop) while reading was also a necessary condition for processing information. However, in later work Baddeley tempered these conclusions, arguing the act of storing information was actually the defining factor for comprehension and that the articulatory loop was only necessary in certain cases. Specifically, Baddeley (1979) reports a study in which participants were asked to repeat either a sequence of six random digits or count continuously up to six while making true/false judgements about sentences. A control condition was also collected, where participants completed only the judgment task with no memory load or articulation required. Baddeley found that the differences in response latency were entirely due to the concurrent memory load and not to the act of having to continuously repeat numbers. Subsequent studies using passage comprehension instead of sentence comprehension confirmed that limiting an individual's ability to store information in the phonological loop did affect comprehension performance but a similar limitation on their ability to sub-vocally rehearse what they are reading did not (Baddeley, 1979). Thus, Baddeley and colleagues concluded that the role the phonological loop plays in comprehension rests almost entirely on its use as an information store and, more specifically, the size of this store, rather than specific control processes such as rehearsal or sub-vocal articulation.

The central executive and reading comprehension

The nature of the central executive, as originally described by Baddeley and Hitch (1974), is to supervise the flow of information between the two subsidiary stores and long-term memory as well as directing attention. Given this role, one would assume that such a mechanism would be necessary to complete any effortful task, so it is surprising that most of the early attempts to describe the processes underlying skilled comprehension within the framework of the multi-component model neglected to empirically specify the contribution of the central executive empirically. The initial reluctance to invoke the central executive (and instead, the over-reliance on the phonological loop) as the link between working memory and reading comprehension was based in an overt admission among researchers that the central executive was too vague or nebulous a concept to be sufficiently useful as an explanatory mechanism, whereas the phonological loop proved more ripe for experimental investigation (Baddeley, 1979; 1986; see also Levy, 1978 for an early critique of the central executive component of the multi-component model). In fact, Baddeley referred to the central executive at one point as the "area of residual ignorance" (1981, p. 21). In the decades following Baddeley and Hitch (1974)'s seminal proposal, however, researchers began tackling the question of what role the central executive may play in reading comprehension— in part because the idea that the phonological loop acts alone and as a mere passive store was unsettling for researchers, especially in light of the presumed progress in research on human memory since the days of the modal model and the idea of memory as simple "boxes" of information.

Daneman, Carpenter, and their colleagues (Daneman and Carpenter, 1980; 1983; Just & Carpenter, 1987; 1992; Daneman & Merikle; 1996) were among the first to implicate the central executive as an integral component of comprehension by examining the relationship from a psychometric perspective. Instead of focusing on the structural components of working memory, Carpenter and colleagues used correlational methods to examine the extent to which performance on working memory tasks can predict individual differences in comprehension ability. Though they opted to use the general term of working memory to refer to domain-general cognitive resources, it was clear that the authors intended this entity to correspond closely to the central executive component of the multi-component model (see Just & Carpenter, 1992 for an explicit comparison of the two concepts). In their paper, Daneman and Carpenter (1980) endeavored to develop a more comprehensive measure of working memory than had been previously employed and outlined several assumptions they used to guide their research. First, they argued that comprehension involves both the storage and processing of information. Readers must store "pragmatic, semantic, and syntactic information from the preceding text and use it in disambiguating, parsing, and integrating the subsequent text" (p. 450). The second assumption was that the system (ostensibly, working memory or the central executive) used to process and store this information was limited, thereby, inducing a tradeoff between the two kinds of activity. If a certain argument or proposition requires heavy processing requirements, the amount of additional information that can be retained should decrease. Conversely, the less complex a passage is, the more information a reader is assumed able to manage. Finally, they argued that there are individual differences in this limitation or "functional working memory capacity," such that some individuals are able to store and process more information, and do so more efficiently, as they read than others. To test these assumptions and investigate the relation between working memory and reading comprehension, Daneman and Carpenter devised a new measure of working memory they called *reading span*. Essentially, reading span required participants to read and process a series of sentences, and at the end of each sentence the participant is required to recall, in sequence, the last words from each of the preceding sentences. The set of sentences increased incrementally until the participant started to make errors. A participant's working memory capacity or "span" was then calculated as the maximum number of sentences on which the participant performed the task perfectly. The authors felt the simultaneous storage and processing aspect of the task distinguished it from earlier measures of short-term memory that relied almost exclusively on storage (e.g., digit span) and posited that the nature of the new measure would make it a better predictor of complex activity such as reading comprehension. This hypothesis proved to be true as the authors found that performance on the reading span task significantly predicted performance on both global measures of comprehension (verbal SAT scores) and specific indices, wherein participants read short narrative passages and were asked to answer fact-based questions and the referent of a pronoun mentioned in the last sentence of a passage. Daneman and Carpenter concluded that an individual's capacity for processing and storing linguistic information directly determined the accuracy and efficiency with which they processed language for meaning. Further, they argued that individual differences in working memory were rooted in the speed and efficiency in which cognitive operations were executed, rather than in variation in the total (or amount of) resources or storage "slots" available. This notion is in stark contrast to the idea of capacity as size, so defined by Baddeley and colleagues, but makes sense if one assumes the central executive to be a facilitator rather than a store.

The introduction of reading span (later expanded to a host of measures coined *complex span tasks*) proved a landmark achievement in understanding the demands of comprehension and to individual difference in working memory research, generally. As will be seen in a later section, the concept of working memory capacity has become an integral (if not the predominant) part of research in human memory and this work owes much of its foundation and methods to Carpenter and colleagues.

The bottleneck hypothesis

The individual difference approach to working memory made clear the necessity and involvement of a limited-capacity resource system in comprehension. However, one of the weaknesses of this work is its dependence on correlational data—indeed, one must be careful not to derive strong causal interpretations using this approach, nor rule out the effect of a shared third variable, such as intelligence.

In an effort to investigate the role of the central executive in reading comprehension using experimental and longitudinal methods, Gathercole, Alloway, and their colleagues (Gathercole, Alloway, Willis & Adams, 2006; Gathercole & Pickering, 2000; 2001; Moulin &

Gathercole, 2007; Alloway, 2006) began examining how the reading process develops across the lifespan—particularly in early childhood and through the primary school years. The results of this research led to the proposal of the "bottleneck hypothesis," which argues that the central executive of working memory acts as a bottleneck for learning, such that poor working memory functioning due to capacity limits prohibits the successful processing and integration of information necessary for completing the classroom activities that students routinely engage in, including acquiring the skills essential for reading comprehension (see also the work of Cain and colleagues, Cain, 2006; Oakhill, Cain, & Yuill, 1998, for a similar proposal that comprehension difficulties can arise due to working memory impairments, even in the absence of lower-level issues such as word reading and vocabulary deficits). Common examples of these activities that form the basis for successful reading development include listening to another speaker, decoding an unfamiliar word while reading sentences and maintaining the meaning of previously decoded text, keeping their place or paying attention while reading a complicated text, or remembering the meaning and content of something they just read for later testing. In each of these scenarios, the student is tasked with processing (often novel) information while also engaging in other demanding cognitive operations. If the student's working memory capacity is too small, the heavy demands imposed by these activities will lead to working memory overload and, thus, frequent task failures. Gathercole and colleagues argue that the cumulative effect of these failures over time leads to pervasive learning difficulties that eventually disturb a student's ability to acquire the foundational knowledge and skills necessary for skilled comprehension.

Support for the bottleneck hypothesis comes from various studies showing that increases in the complexity of classroom activities leads to decreases in success and this drop in performance is directly related to working memory capacity. In a study by Gathercole, Lamont, and Alloway (2006), school-age children completed measures of both verbal working memory and literacy. The authors found that as the demands on working memory increased, children identified as having low working memory abilities were found to make errors more frequently. Specifically, high rates of failures were observed in activities that required children to follow complex instructions or perform tasks with a complex hierarchical structure. In other words, these activities required the coordination of a number of pieces of information that some children did not have the innate capacity to manage. In another study, Gathercole, Brown, and Pickering (2003) examined the association between working memory function and academic achievement over the first three years of school. Specifically, they found that a child's verbal working memory ability assessed at the start of formal schooling was predictive of their performance on reading and language assessments in later grade levels. Simply put, poor working memory function early on directly constrained the capacity for children to learn the complex skills and knowledge required for academic success later in school.

Though the bottleneck hypothesis of Gathercole and colleagues is a departure from the resource sharing theory of Daneman and Carpenter (1980) in the sense that it invokes Baddeley's original ideas of capacity as the amount of information that can be processed, it nonetheless corresponds to the idea that the central executive is the focal point of working memory functioning and the primary mechanism for completing effortful cognitive operations such as comprehension.

The embedded process model of working memory

One of the most common critiques of the multicomponent model of working memory is its inability to account for the full range of sensory information that individuals experience (often simultaneously) beyond simply aural and visual (e.g., tactile, olfactory, gustatory, etc.). While the model has been updated with added components to account for contextually rich memory representations (see the introduction of the episodic buffer, Baddeley, 2000), it is difficult to imagine all of these stimuli and feature types being parsed into several different, discrete stores. Instead of concluding that the many different types of features each have their own modules, researchers as early as Atkinson and Shiffrin (1968) have suggested that short-term and long-term storage does not require that the two stores necessarily involve different structures but rather, it is possible that the short-term store be considered simply as temporary activation of some portion of the long-term store (see also Anderson, 1983). Similarly, van Dijk and Kintsch (1983), in proposing how thematic elements and topics of a text may influence comprehension without being in consciousness, proposed a central processing unit consisting of an "active, conscious but strictly capacity-limited core, and a boundary area containing the control system where memory structures affect processing in the core without being directly conscious and without being subject to capacity limitations. Knowledge structures beyond that boundary…can participate in discourse processing only if they are successfully retrieved" (p. 351).

The idea that information and "memories" exist on a continuum of activation that is not modality specific was further developed by Nelson Cowan (1988; see also Cowan, 1995, 1999) and his embedded process model of working memory. In this model, maintenance, rehearsal, and retrieval processes work upon activated traces of long-term memory, rather than separate representations held in domain-specific storage structures like those of the multicomponent model. Instead, memory is considered a unitary construct but with three distinct levels of activation: (a) inactive representations of long-term memory, (b) an activated subset of long-term memory considered the short-term store, and (c) a subset of the activated representations in the short-term store available to conscious awareness labeled the focus of attention (see Figure 3).

The short-term store and the focus of attention collectively make up working memory and are subject to specific limitations. In the short-term store, any number of long-term memory representations can be activated at any one time but the level of activation decreases over time, typically in the realm of 10–20 s unless reactivated. In contrast, the focus of attention is limited by its capacity rather than by time. Specifically, Cowan argues individuals can maintain only 3–5 unrelated items in the focus at once, all of which are susceptible to displacement back into the short-term store by new information (Cowan, 2001). Cowan also posits that managing the memory system is a central executive, akin to the multicomponent model, that is responsible for all control processes. The focus of attention, in particular, is regulated (in part) by the central executive while also being controlled by involuntary mechanisms (such as distraction or attentional capture). Beyond the coding of new information, the central executive is also integral to the activation of stored representations in long-term memory.

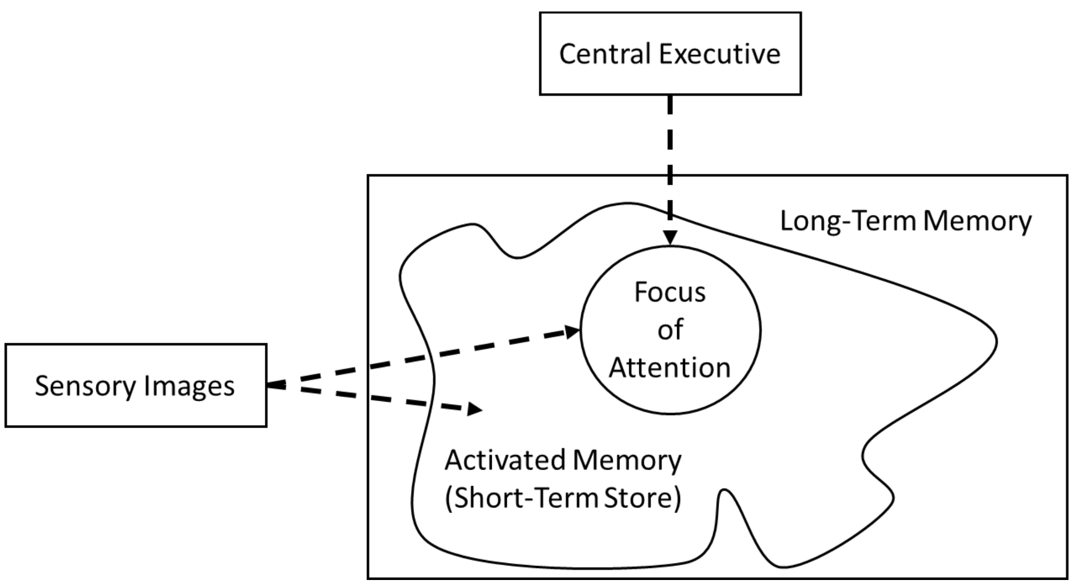


Figure 3. Schematic representation of the embedded process model of working memory. Adapted from Cowan (2001).

The embedded process model of working memory and reading comprehension

In contrast to the multicomponent model, little research has been done to directly apply the embedded process model to the problem of reading comprehension. This may be due in part to the generalized nature of the model; that is, Cowan has provided a framework—replete with limits and boundary conditions—for information processing with the assumption that neither the type of material to be processed nor the purpose for processing changes how the system operates. The act of reading and understanding text is expected to function under the same conditions as experimental list-learning paradigms. Individuals direct the focus of their attention to a particular source of stimuli, which, in turn, activates existing long-term memory representations. These new and old representations are then recombined to form a new, meaningful pattern of features or concept. Activated memory representations are assumed to take not only the form of individual words but syntactic, thematic, and semantic information as well. In fact, Cowan (2014; 2015) suggests that the focus of attention holds special significance compared with the activated subset of information represented in working memory outside of the focus in that it serves as a "workshop," whereby new concepts are formed by binding together information that is currently in focus to form new information. For example, the co-occurrence of *big* + *striped* + *cat* in the focus of attention forms the concept of a *tiger*. Evidence for the associative nature of the focus of attention comes from Cowan, Donnell, and Saults (2013), in which participants were provided lists of three, six, or nine items at once and were asked to indicate which item in the list they considered "most-interesting." Later, participants were given a surprise test on whether pairs of words came from serial positions in the same list or from nearby serial positions in different lists. The authors found that performance was much better for lists of three items than for the longer list-lengths, which were not much better than chance, and argued this was because nearby items from the same list would have coexisted in the focus of attention much more often for lists of three items (within the presumed 3-5 item capacity of the focus of attention) and less often for longer lists of six or nine items (for a related discussion of temporal contiguity effects in free recall paradigms, see also Kahana, 1996). The authors concluded that these results support the idea that the linking together of elements to form new concepts is primarily handled within the focus of attention.

Cowan (2001) calls the linking together of elements, *chunks,* and uses this concept much in the same way that Kintsch and colleagues did; namely, that chunks represent related, conceptual elements. Cowan argues that the focus of attention is not necessarily limited by items but by these chunks, such that more information can be maintained if chunked effectively. The aforementioned example of *tiger* being derived from the association of the features *big + striped + cat* provides a simple instance of how multiple elements can be processed into a single concept to reduce storage demands. Another common example of how chunks operate is the letter string: I R S F B I C I A. On the surface, this letter string should exceed the stated capacity of the focus of attention; however, one can use prior knowledge to chunk the letters into more meaningful units: IRS, FBI, and CIA. The storage demands of the letter string now equate to only three chunks compared with the original nine. Cowan uses this line of reasoning to predict how the embedded process model handles the act of comprehension (Cowan, 2014; 2015). Limits in the focus of attention prohibit a reader’s ability to commit every minute detail of what they are reading to memory, so to accommodate, they retain the semantic gist elements of the text and mostly ignore the syntactical or orthographic details. As the reader proceeds through the text, they essentially chunk these gist concepts and ideas into separate parts or propositions until they can integrate them into a united semantic structure, all the while disregarding the verbatim information because working memory cannot hold it (for evidence of this idea in language comprehension, see Ferreira, Bailey, & Ferraro, 2002). Finally, Cowan argues that as an individual develops cognitively throughout childhood, the amount of items that can be processed and associated in the focus of attention increases; thereby, equating to a concomitant increase in the complexity of ideas that can be comprehended (Cowan, 2014).

Individual differences in working memory capacity

Rather than attempting to formulate a specific model of working memory and then align components of this model to aspects of reading comprehension, many researchers have increasingly begun to use an alternative method for investigating the working memory and reading comprehension relationship—namely, by using a differential approach to study individual differences and emphasize the function of working memory rather than its structure (Engle & Kane, 2003). Beginning with the aforementioned work of Daneman and Carpenter (1980), a substantial amount of research has been conducted to develop measures of working memory capacity and employ them as predictors of performance on a host of cognitive activities, including reading comprehension. Though various methods have been used in this approach and no single model of working memory has been adopted, it is widely assumed that these measures of working memory capacity (termed working memory or complex span tasks) are indicators of the actions of the central executive (Daneman & Carpenter, 1980; Just & Carpenter, 1992; Cain, 2006; Baddeley, 1992; Kane, Bleckley, Conway, & Engle, 2001). Moreover, the defining feature of the psychometric approach is that tasks purporting to measure working memory must involve both the storage and processing of information and that individual differences arise in the ability to handle this tradeoff (Daneman & Carpenter, 1980; Turner & Engle, 1989). Support for these assumptions comes from work showing that complex span tasks are significantly better predictors of higher-order cognitive performance compared with traditional short-term memory simple span tasks (Daneman & Carpenter, 1980; Daneman & Merikle, 1996; La Pointe & Engle, 1990; Kane, Bleckley, Conway, & Engle, 2001; although see Unsworth & Engle, 2007a for a discussion of list-length effects and the predictive power of simple span tasks). Historically, researchers have argued that the predictive discrepancy between the two types of tasks rest squarely on the dual-task nature of complex span task. In a sense, complex span measures more accurately mimic the processes required during problem-solving, reasoning, and comprehension—storage of information alone is not sufficient for success, there must be concurrent processing and integration of information.

Controlled attention view of working memory capacity

The most prominent theory of why complex span measures are such great predictors of higher-order cognition is the controlled attention view (also called executive attention or attention control) of Randall Engle, Michael Kane, and colleagues (Engle, Tuholski, Laughlin, & Conway, 2001; Engle, Kane, & Tuholski; 1999; Kane, Bleckley, Conway, & Engle, 2001; Engle & Kane, 2003). In this view, the correlation between working memory span measures and higher-order cognitive capabilities, such as intelligence or comprehension, is driven almost exclusively by the general ability to sustain an active memory representation in the face of attention shifts (due to interference or distraction). Thus, individual differences in working memory capacity are not really about memory storage, but about controlling attention such that goal-relevant information remains highly active, and is accessible under conditions of competition. Converging evidence for a general mechanism comes from a variety of studies showing that the specific concurrent-processing task used to measure working memory capacity has little effect on its ability to predict complex cognition (Turner & Engle, 1989; Kane, Bleckley, Conway, & Engle, 2001; Engle & Kane, 2003). This finding is in contrast to Daneman and Carpenter (1980; 1983)'s initial proposal that working memory capacity limits were tied to the specific processing demands of the concurrent task (in their case, reading sentences). They argued that good readers did well on the reading span task because they could devote less resources to processing during reading and more to storage than poor readers, but left open the idea that good and poor readers could have equivalent capacities during other, non-reading tasks. Turner and Engle (1989), however, disputed this idea and argued that working memory capacity reflected a much more general capability that is domain-free and independent of any one type of processing task. To support this notion, Turner and Engle modified the reading span measure to incorporate mathematical processing (called *operation span*) and found that it still significantly predicted performance on measures of reading comprehension like the Nelson-Denny Reading Comprehension Test and the SAT, suggesting that individuals who performed poorly on the reading span measure did not do so simply because they were poor readers but because they had a more general working memory deficit (see also La Pointe & Engle, 1990).

Specific support for the controlled attention view of working memory capacity comes from both latent-variable and experimental analyses. Engle, Tuholski, Laughlin, and Conway (1999) had a large sample of participants complete a battery of tests, including complex span measures (each differing in the nature of the concurrent processing task), three traditional simple span measures of short-term memory, and two established tests of general fluid intelligence, the Ravens Progressive Matrices and the Cattell Culture Fair Test. Exploratory factor analysis and structural equation modeling were performed on the results. The authors found that each batch of tests formed a distinct latent factor (i.e., a WMC factor, a STM factor, and a gF factor). Further, the three factors were strongly related to each other, however, once the shared variance between the WMC and STM factors was removed, the correlation between the WMC factor and gF remained strong (r = .5) while the residual STM factor showed no significant relation to gF. The authors argued that the shared variance between WMC and STM reflected storage, while the residual of the WMC factor that was highly correlated to intelligence represented the controlled attention component. In a later experimental study, Kane, Bleckley, Conway, and Engle (2001) had participants complete complex span tasks and then divided these participants into high- and low-working memory capacity span groups based on their performance. These two subgroups were then asked to complete a visual saccade task. The visual saccade task contains two variants: *pro*-saccade trials and *anti*-saccade trials. During pro-saccade trials, participants keep their eyes upon a fixation point in the center of a computer screen until a flashing cue appears in their periphery on either the left or right side of the screen. Once the cue appears, the participant is required to direct their eyes to the direction of the cue to see a target stimulus. The target stimulus is either *B*, *P*, or *R* and flashes very quickly before being masked by the number *8*. The participant must then make a response indicating which target stimulus was presented before moving back to center fixation and beginning the next trial. During anti-saccade trials, participants are required to move their eyes in the *opposite* direction of the flashing cue to see the target stimulus. Kane et al. (2001) found that the two sub-groups performed equally well in pro-saccade conditions; however, low-span participants were much slower and less accurate than high-span participants in the anti-saccade conditions. A follow-up experiment using eye-tracking found that low-span participants were more likely to be captured by the flashing cue and slower to correct their saccade errors. The authors concluded that anti-saccade tasks required active maintenance of the task-goal (i.e., look opposite way of cue) and suppression of the prepotent response to look at the cue. If an individual's ability to flexibly control attention is lacking, as is assumed for low-span participants, they are more likely to be distracted by irrelevant stimuli and "forget" what they are supposed to be doing, thereby, leading to a higher rate of task failure.

Controlled attention and reading comprehension

Without question, the bulk of research on individual differences in working memory and complex cognition has revolved around intelligence; however, interest in testing the controlled retrieval view within the domain of reading comprehension is growing. Specifically, McVay and Kane (2012; see also Unsworth & McMillan, 2013) recently used latent-variable analysis to investigate whether mind-wandering mediated the relationship between working memory capacity and reading comprehension. Mind-wandering is the phenomenon in which an individual has trouble maintaining attention during an ongoing task or activity due to intrusive task-unrelated thoughts (or TUTs)—essentially, TUTs cause lapses of attention, which, in turn, lead to a decrease in task performance (Kane, Brown, McVay, Silvia, Myin-Germeys, & Kwapil, 2007; McVay, Kane, & Kwapil, 2009). Naturally, the controlled attention view argues that mind-wandering is the result of low working memory capacity causing an inability to maintain on-task thought (McVay & Kane, 2010). McVay and Kane argued that if controlled attention is the primary mechanism underlying the relation between working memory and higher-order cognition, such as reading comprehension, then one should expect mind-wandering to act as a mediator. Specifically, individuals with lower working memory capacity should have greater comprehension deficits, in part because they are less able to maintain on-task thought and more susceptible to distraction. To test this idea, McVay and Kane had a large sample of participants (*N* = 242) complete multiple indices of working memory capacity, reading comprehension, and attention control tasks. To measure mind-wandering empirically, the authors briefly interrupted participants during two of the reading comprehension tasks and two of the attention control tasks to ask if their immediately preceding thoughts could be classified as on-task or off-task. After conducting structural equation analysis on the results, the authors found that TUT rates remained relatively stable across tasks and that this general susceptibility to off-task thought was detrimental to performance in those tasks. They also found that working memory capacity negatively predicted TUT rates and positively predicted reading comprehension and, moreover, that the shared variance between working memory capacity and attention control appeared to underlie the working memory capacity and reading comprehension association. Finally, TUTs was found to mediate the correlation between working memory capacity and reading comprehension, suggesting that control over thought content is a necessary condition for successful comprehension and is one of the means through which working memory influences comprehension ability. A similar study by McVay and Kane (2010), found that when higher working memory capacity participants report TUTs, they perform as poorly as low working memory capacity individuals.

Good text comprehension requires, among other things, decoding unfamiliar words, the resolution of pronouns, memory for facts, and inferential reasoning. The work of McVay, Kane, and colleagues offers a compelling argument that working memory is important for reading comprehension such that working memory resources are necessary to maintain focus during reading so that these actions can be successfully undertaken. Indeed, they have shown that the presence of TUTs displaces the task goal of comprehending reading material and this has a deleterious effect on a reader's ability to process relevant details and build a coherent mental model of the text necessary for adequate comprehension. One can imagine that if a child learning to read does not have the capacity to manage distraction and interference, whether internally (e.g., intrusive thoughts or motivation issues) or externally (e.g., noise or pictures), their ability to learn the skills necessary to compete at their grade level or beyond will be severely limited.

Long-term memory retrieval: The missing link between working memory and reading comprehension?

Much of the recent research using a psychometric approach to explain the association between working memory and reading comprehension skills has emphasized the importance of attention or, more specifically, the ability of an individual to maintain focus on and process several pieces of information at once. As the work of Engle, Kane, and others reviewed above indicate, there is support for the idea that individuals with greater attentional capacity (or control of attention) will have less difficulty processing and understanding text. Indeed, the prevailing view among both working memory and individual difference researchers is that working memory capacity equates to the ability to maintain the focus of attention on a subset of information or in an ongoing task. If this information or task-goal is displaced from the focus of attention because of capacity-overload, lapses in attention, or interference, performance on a task declines and failure becomes more likely. However, the discussion of working memory as a construct primarily concerned with the control of attention and the active maintenance of information leaves open the possibility that the ability to retrieve information that has been lost from focus of attention may be a distinguishable skill as well. In other words, if individual differences exist in the control of attention, they may also exist in retrieval and, more importantly, these differences may be just as important for complex cognitive functioning. Further, it allows for the possibility that an individual could be labeled as having poor working memory function despite having the same "capacity" as his or her peers in terms of the amount of information that can be held in immediate memory or their ability to flexibly allocate attention. That is, although the individual could maintain the same amount of information, they could lack the sufficient ability retrieve this information back into mind if lost or "forgotten." Anecdotally, we know that even after information is displaced from immediate memory, this information is not always completely lost but may simply require a search of long-term memory for its restoration. Indeed, research has shown that individual differences in long-term memory retrieval ability may be due primarily to issues with accessibility of a specific memory rather than availability (Unsworth, Spillers, and Brewer, 2012; Tulving & Pearlstone, 1966). The question becomes then, why have researchers not examined the contribution *retrieval from long-term memory* may have in accounting for the relationship between working memory and higher-order cognition?

Interestingly enough, when proposing how to reduce the cognitive demands on students in classroom activities and, ostensibly, account for deficits in attention control, many researchers suggest teaching approaches that actually appear to reduce the need for students to have to initiate an effortful search of long-term memory. Some of these recommendations, for example, include breaking tasks and instructions down into smaller steps, simplifying the linguistic and syntactical structures of to-be-learned material, increasing the familiarity of what is being read, repeating information multiple times, encouraging the students to ask if they have forgotten what they should be doing, and providing external reminders or cues in the classroom environment (e.g., wall charts) so that task demands are more manageable (Alloway, 2006; Holmes, 2012). Notably, the common thread among all of these recommendations is the idea that the root cause of performance failures in a given task is due to capacity overload in working memory and that educators should, therefore, provide externally provided support so that if information is displaced (or "forgotten") from short-term memory, a self-initiated search of long-term memory to restore that information is not required. This implies that long-term memory retrieval is an integral part of successfully handling classroom task demands.

This is not an altogether new idea. Kintsch (1982) in discussing the limitations of the Kintsch and van Dijk (1979) Construction-Integration model calls for more investigation of the role retrieval may play in text comprehension, especially given its central role in memory theory, generally. Similarly, the inclusion of the short-term activated memory feature of Cowan (1988; 1995)'s embedded process model is an overt attempt to incorporate long-term memory retrieval into working memory functioning to resolve the paradox of how people perform quite well in complex, everyday environments, despite working memory being limited to a few ideas or items at once. In a more recent study examining the relationship between working memory and mathematical problem solving, Swanson and Beebe-Frankenberger (2004) suggested that working memory provides a resource that allows the learner to integrate information retrieved from long-term memory with information currently being held in focus and that the influence of working memory performance on problem solving is related to one's ability to accurately access information from long-term memory to solve the problem (see also Swanson & Saez, 2003). With reference to tasks specific to reading, it is not difficult to conceive of a similar situation arising in which retrieval from long-term memory is a prerequisite for success. Imagine, for example, an individual is asked to read a passage on an unfamiliar topic that is several paragraphs long and then answer questions about what they have read. To complete this task, the individual is required to do many things. First, they must keep the task goal in mind ("understand what is being read well enough to correctly answer questions"). Second, they must process and synthesize each line of text in order to form a coherent representation (or situation model) of what is being discussed. As the reader progresses through the material, they must store and constantly update the ideas that are being presented as newer information is being processed and, furthermore, try to connect this new information with their own prior knowledge. Depending on the complexity and length of the text and the novelty of the content, an individual could become quickly overwhelmed by the amount of information they are required to keep in mind all at once and, indeed, it is likely certain pieces of information will be displaced from their immediate memory. Therefore, in order to answer the questions that may be asked of them successfully, it is important that a reader can initiate a search process of long-term memory to restore this information. In the event the reader has difficulty successfully retrieving the required information, task failure is expected no matter how large their attentional capacity.

Long-term working memory

The first detailed attempt to incorporate the process of retrieval from long-term memory into models of working memory comes from Ericsson and Kintsch (1995). The development of their long-term working memory model originated out of the dilemma of how individuals are supposed to handle the complexity of most problems using only a limited capacity, short-term mechanism. They did not believe it was possible to explain how memory was used in tasks such as solving complex math problems or reading comprehension within the standard working-memory framework. Building upon the findings gleaned from research on chunking (Miller, 1956; Cowan, 1995) and skilled memory theory (Chase & Ericsson, 1982), the long-term working memory model proposed that individuals rely on specific control processes to encode information in long-term memory in readily retrievable form. Specifically, long-term working memory is conceived as a subset of long-term memory that is directly retrievable via cues that are stored in short-term working memory (or the focus of attention). Because the desired information is already encoded in long-term memory, once a cue enters the contents of the short-term working memory system, retrieval is fast and automatic. Thus, the contents of short-term memory automatically create or activate long-term working memory. Essentially, a network of permanently encoded memory representations is activated by a short-term cue. Given the symbiotic nature of short-term and long-term memory in this model, Ericsson and Kintsch acknowledge that long-term working memory is not argued to be a replacement for short-term or working memory, instead, long-term working memory is necessary only in particular domains to meet specific storage and retrieval demands imposed by a given activity. In other words, short-term or working memory may be sufficient for simple list-learning paradigms employed in the laboratory, but for more complex activities such as playing chess, long-term working memory will be required for maximum performance.

To provide support for their model, Ericsson and Kintsch (1995; see also Kintsch, Patel, & Ericsson, 1999) applied the long-term working memory model to text comprehension. In their paper, they viewed text comprehension as expert performance—that is, skilled comprehenders are experts at reading in the same way a professional athlete is at their chosen sport. Using the previously discussed Construction-Integration model of text comprehension (Kintsch, 1988, 2004) as a guide, the authors argued comprehension results in the formation of memory traces (called memory *nodes*) in long-term memory from propositions derived directly from the text. Additionally, these memory nodes are linked in a complex network. As a reader proceeds through text, new text nodes (words, sentences, concepts) entering short-term working memory are linked to older text nodes in long-term memory through the process of comprehension. The strength of these connections is dependent on the nature of the text and the specific comprehension strategy that a reader employs. According to the C-I theory, if a text is comprehended properly, a coherent mental representation has been generated and links between mental nodes are stable and strong, permitting automatic retrieval. Although not expressly discussed by Ericsson and Kintsch (1995), the long-term working memory framework predicts that comprehension deficits are due less to the capacity of short-term working memory and more due to a lack of appropriate comprehension strategies, prior knowledge, and language skills.

A new perspective: The dual-component model of working memory capacity

While the long-term working memory model was a significant advancement in how we think about the role of long-term memory in working memory functioning, Ericsson and Kintsch (1995) were quick to clarify that recruitment of a long-term working memory system is restricted to well-practiced tasks and familiar knowledge domains. In their view, retrieval in the long-term working memory system is considered to be automatic as the path between retrieval cues contained in short-term memory and the associated network of memory representations in long-term memory have already been well-established. As discussed earlier, however, there is still reason to believe long-term memory retrieval plays an integral role in the working memory system, more generally—that is, even in situations in which individuals are not experts and are in the midst of the actual learning process. In fact, recent correlational work has suggested that measures of working memory are highly related to measures of long-term memory and, further, can explain variation in individual differences in working memory capacity beyond what is explained by measures of attention control or capacity (Unsworth & Engle, 2007b; Unsworth, 2010; Unsworth, Brewer, & Spillers, 2009; Mogle, Lovett, Stawski, & Sliwinski, 2008).

To account for these new findings, Unsworth and colleagues (e.g., Unsworth & Engle, 2007a; 2007b; Unsworth & Spillers, 2010) proposed the dual-component model of working memory capacity to explain individual differences in working memory. As the name implies, the dual-component model incorporates two distinct components or control processes that act on both short-term and long-term memory stores: 1) active maintenance of information in primary (or short-term) memory and 2) controlled retrieval of information from secondary (or long-term) memory. Primary memory is argued to reflect both the number of items that can be distinctly maintained in focus and attention control processes that prevent attentional capture or mind-wandering by continually allocating attentional resources to maintaining those items. These "items" or representations could include things such as the task-goal, encoding strategies, or actual units of information such as words from a list or text. As long as attention is allocated to these representations, they will be actively maintained in primary memory. If attention is removed, however, due to internal or external distraction or to process new information, these representations will no longer be active and will need to be retrieved from secondary memory. Secondary memory reflects the need to retrieve items that could not be maintained in primary memory (either due to distraction or overloaded capacity limits) as well as the need to retrieve other task-relevant information from long-term memory, such as prior knowledge or schemas. Importantly, retrieval from secondary memory is cue-dependent. The extent to which representations can be retrieved will be dependent on encoding abilities, the ability to reinstate encoding context at retrieval, and, more importantly, the ability to focus the search on target items and exclude interfering items (Spillers & Unsworth, 2011).

Support for the dual-component model comes from latent-variable analyses showing that performance on both attention control and long-term memory retrieval tasks jointly predict performance on complex working memory span tasks and measures of higher-order cognition. Specifically, Unsworth and Spillers (2010) conducted structural equation models examining the extent to which attention control abilities and long-term memory retrieval abilities accounted for variation in working memory capacity and its relation to fluid intelligence. In their study, participants completed various measures of attention control, long-term memory retrieval, working memory capacity, and fluid intelligence. Confirmatory factor analyses conducted on the data revealed that each construct was best represented by separate, yet correlated factors. Structural equation modeling revealed that attention control and long-term memory retrieval accounted for unique variance in working memory capacity and, further, that a significant part of the shared variance between working memory capacity and fluid intelligence was due to both attention control and long-term memory retrieval abilities. In a follow-up study, Unsworth, Fukuda, Awh, and Vogel (2014) tested the prediction of the dual-component model that there should be individual subgroups in performance on one or both of these components. That is, if each component is distinct and not controlled by a single, unitary supervisory system, one would expect it to be possible that an individual can adequately maintain information in the focus of attention in the face of distraction but not be adept at successfully retrieving this information if lost from focus, and vice-versa. Unsworth et al. (2014) had participants again complete various measures of attention control, long-term memory retrieval, working memory capacity, and intelligence. Importantly, the authors also distinguished between measures of attention capacity and those of control. Aside from replicating the findings of Unsworth and Spillers (2010), latent variable analysis also revealed that both storage of information and the ability to control attention were uniquely related to measures of working memory capacity and intelligence. More importantly, cluster analyses performed on the data revealed that five distinct groups of participants could be defined based on task performance across each of the measures. Specifically, some participants had clear deficits in the number of items they could maintain in the focus of attention but little trouble in controlling attention in the face of irrelevant distractors, while another group of participants demonstrated the exact opposite pattern of deficits. Additionally, one group of participants performed well in attention tasks but had specific issues with retrieval from long-term memory. Finally, some participants had deficits in attention control, capacity, *and* retrieval, while another group showed high-levels of performance in every domain. The authors concluded that the results provide convergent evidence to the multifaceted nature of working memory and that working memory deficits can arise for a number of reasons.

The dual-component model and reading comprehension

With this framework in mind, we can predict how the dual-component model would account for performance in a task like reading a passage on an unfamiliar topic and answering questions about what has been read. The task goals ("understand what is being read well enough to correctly answer questions"), words in the text, and conceptual information are all fighting for attention in a limited capacity store (e.g., attentional focus) and depending on a reader's ability to maintain this information all at once and the complexity of each piece of information, some bits may fall out of focus, necessitating retrieval of this lost information from long-term memory. Therefore, variation in either (or both) component process would affect the likelihood of success. Similarly, in many reading comprehension assessments, individuals are required to read a text in full before being asked to summarize what they have just read or answer various questions concerning its meaning, the characters, plot points, etc. Certainly we would not expect an individual to maintain every detail all at once while reading, therefore, it is assumed that they will need to rely on a strategic search process to recall the necessary information back into mind. Failure in the ability to do so efficiently will lead to slower response latencies, more flipping through pages to find what they are looking for (assuming they can remember where in the text their answer is likely found), or even task abandonment altogether.

Though not a direct test of the influence that retrieval ability may have on the relationship between working memory and reading comprehension, Unsworth, McMillan, Spillers, and

Brewer (2013) had participants complete a battery of cognitive ability measures (including working memory capacity and long-term memory tasks) and then log memory failures in a diary over the course of one week. Participants were asked to include any incidents in which they failed to retrieve a particular memory or piece of information (e.g., someone's name, where keys are located, factual information for a test, etc.). The authors found that not only were memory failures related to working memory capacity but they were also related to verbal SAT performance (including passage comprehension). Interestingly, certain memory failures were more consistently related to working memory capacity and verbal SAT performance than others. Specifically, those having to do with educational contexts correlated more strongly than more general failures, such as forgetting someone's name. The authors argued that the current results provide evidence that the ability to retrieve target information is a major source of individual differences in scholastic abilities, as measured by the SAT.

A new proposal: Applying the dual-component model to reading comprehension

Though the initial endeavor by Unsworth et al. (2013) to apply the dual-component model of working memory to scholastic abilities is promising, it underscores the need for more research examining the possible influence that controlled retrieval may have on the development of *specific* academic skills (e.g., reading comprehension) beyond simply correlating single measures of long-term memory with global measures of scholastic ability. Indeed, the purpose of the current study is to conduct a more targeted and comprehensive approach to investigating the nature of retrieval in the working memory-reading comprehension relationship, much like those done with working memory and intelligence. Specifically, latent variable analysis will be used in an effort to examine the influence that the specific components of working memory may have on performance on commonly employed assessments of reading comprehension.

CHAPTER 2

THE CURRENT STUDY

Working memory is the retention of a small amount of information in a readily accessible form, which facilitates problem-solving, learning, reasoning, and comprehension. As should be evident from the above review, working memory includes not only memory itself but also the executive control skills used to manage the information in working memory (see the *working-with-memory* model of Moscovitch, 1992). Further, working memory is assumed to be limited both in terms of the amount of information that can be processed and the efficiency in which it does so—although there is continued debate about the nature of these limitations and what they mean for information processing. Given these limitations, many researchers have begun to investigate the somewhat paradoxical role that long-term memory plays in working memory functioning. Much of the recent research examining individual differences in working memory has found that the ability for an individual to quickly and accurately retrieve information from long-term memory is just as important for complex cognition (such as intelligence or mathematical problem-solving) as the ability to control attention in the face of distraction.

With respect to reading comprehension, it is clear that working memory is *the* primary mechanism responsible for processing text and creating permanent and coherent memory representations about what has been read. The specific components of working memory that influence reading comprehension performance, however, remains an open question. By and large, most memory researchers view working memory essentially as an attentional system. Deficits in reading comprehension, therefore, have been primarily explained by a reader's ability (or inability) to maintain focus on, or store, what they are reading long enough to process this information and connect it for long-term use. Sometimes readers struggle with information overload and do not have the capacity to manage the complexity of what they are reading, while other readers struggle with issues of distraction and interference leaving them susceptible to lapses of attention. Once information is displaced from short-term memory or the focus of attention because of capacity-overload, distraction, or interference, the question then becomes, *what happens next*? Certainly, it is possible that restoration of what has been read back into the focus of attention via a long-term retrieval process is as important as keeping that information in focus in the first place.

Whether retrieval of information is actually as important as control of attention for reading comprehension performance has not been empirically tested. Indeed, no study has directly investigated the possible role that controlled retrieval from long-term memory may play in predicting performance on measures of reading comprehension. Therefore, the purpose of the current study is to implement a large-scale ($N > 200$) correlational study using latent-variable analysis and examine the specific control processes that may underlie the relationship between working memory and reading comprehension.

Planned Analyses

A combination of confirmatory factor analysis and structural equation modeling will be used answer the primary question of interest—namely, how is working memory related to reading comprehension performance? Specifically, confirmatory factor analysis will be used to examine the structure of the underlying data, followed by structural equation modeling to test the relationships between latent variables and my central hypothesis. Finally, a sub-group analysis will be conducted in an effort to parse patterns of participant performance into distinct clusters or “groups.” Such an analysis will be important for revealing any nuances in working memory functioning that may differentially affect reading comprehension performance.

1. Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) is a statistical technique for examining covariance among several different indicators or measures, with the expectation that there is a latent factor or construct creating an association between one or more of these indicators.

In this study, four latent factors will be derived and the relationships between these constructs examined. These factors will include working memory capacity (WMC), attention control (AC), long-term memory retrieval (LTM), and reading comprehension performance (RC). Three indicators are to be used to measure each latent variable. While previous studies (e.g., Unsworth et al., 2013; Daneman & Carpenter, 1980; Gathercole, Brown, & Pickering, 2003; Turner & Engle, 1989) have relied on a single task to represent a given construct, the use of multiple indicators reduces the possibility that results could be biased due to task-specific variance and also decreases the likelihood of technical problems during the model estimation. It is recommended that at least two indicators be used for each respective factor (Kline, 2016). More information regarding the indicators and the specific factors for which they are being used is discussed in the next section.

As noted in the above review, previous research has found: 1) that WMC and RC are related, 2) that AC and LTM are related to WMC, and 3) that AC is related to RC. The purpose of the proposed CFA is to see if these relations hold true with the assessments of each construct examined in the present study, so that structural equation modeling can be used to determine causal relations among the factors. Secondary goals for conducting the CFA include confirming: a) that the selected indicators are appropriate for their associated factors and b) that there are positive correlations between the LTM factor and the other factors before continuing to the structural equation modeling analysis. A mock CFA model depicting the hypothesized grouping of indicators with their respective latent factors and proposed relations between factors can be seen in Figure 4.

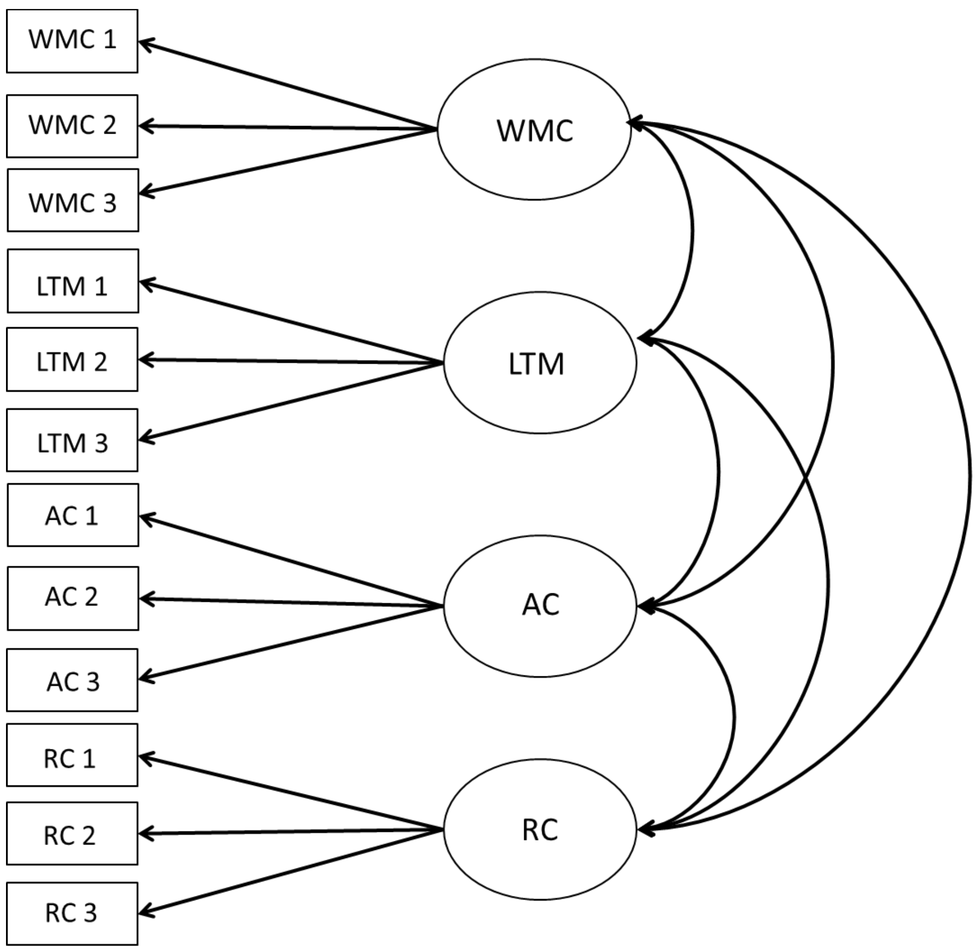


Figure 4. Mock diagram of hypothesized confirmatory factor analysis with four latent constructs (circles) and three associated indicators (rectangles) each. Factors: working memory capacity (WMC), long-term memory (LTM), attention control (AC), and reading comprehension (RC).

2. Structural Equation Modeling

The type of structural model to be employed in the current study is considered a hybrid model, consisting of both observed and latent variables. Hybrid structural models are those that include both measurement and structural model components—essentially, a CFA and path analysis combined. In the present study, once a measurement model has been established using CFA, a structural model detailing causal relationships between the various factors will be developed based on the predictions of the dual-component model of working memory. Structural equation modeling (SEM) is a useful tool for answering the primary question of interest because it allows one to test a hypothesized model against an observed covariance matrix—that is, it allows one to examine the relationships between variables that cannot be directly observed. What distinguishes this technique from CFA is that SEM also allows researchers to test causal relations between latent factors (variable X → variable Y or vice-versa). Moreover, it is important to note that the researcher is responsible for specifying the model to be tested. It is not enough for a model to best "fit" the data but do so under the auspices of a defined hypothesis about how the constructs should relate to one another. It may be the case that a researcher expects one variable to predict another but, in reality, this specification may be inconsistent with the data. Therefore, modifications must be made to not only the model specification but to one's theory as well. It also important that the specific tasks or measures chosen to form a particular construct have been adequately vetted and are theoretically appropriate for inclusion as an indicator in the model.

To answer the specific question of whether the AC and LTM factors explain the relationship between WMC and RC, a model will be tested to investigate the likelihood that there are indirect effects of the AC and LTM factors on the relationship between WMC and RC. As discussed, the dual-component model of working memory argues that both AC and LTM abilities underlie individual differences in WMC functioning. As WMC and RC are highly correlated, it follows from the dual-component model that both AC and LTM should be uniquely related to reading comprehension performance as well. Further, AC and LTM abilities should mediate the relationship between WMC and RC, such that the direct effect should no longer be significant once indirect effects of AC and LTM factors have been accounted for (see Figure 5 for a mock diagram of this hypothesized model).

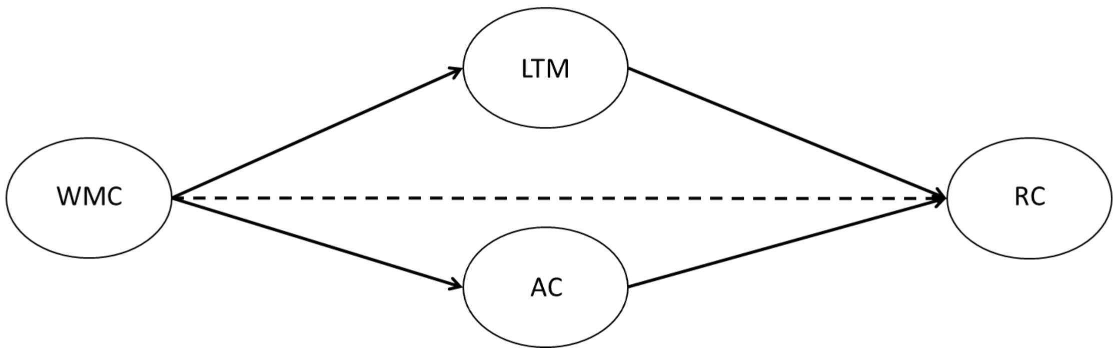


Figure 5. Mock diagram of hypothesized structural equation model indicating the nature of relationships between latent constructs. Specifically, we predict indirect effects of attention control (AC) and long-term memory retrieval (LTM) factors on the relationship between working memory capacity (WMC) and reading comprehension (RC).

3. Sub-group Analysis

Sub-group analysis is an unofficial term to describe categorizing the data according to specific patterns in participant performance or clustering groups of participants based off shared performance characteristics. Essentially, clusters are determined by minimizing within-group differences and maximizing between-groups differences. Groups are formed where individuals in the group are very similar to one another but unlike individuals in other groups (see Everitt, Landau, & Leese, 2001; Kaufman & Rousseeuw, 2005; Unsworth, 2009).

If our assumptions are true, that both AC and LTM can independently affect performance on measures of reading comprehension, then it is possible that readers can fall into distinct groups based on their respective AC and LTM abilities. Specifically, one would expect four different patterns of performance: 1) those who perform well on both AC and LTM tasks, 2) those who perform well on AC tasks but not LTM, 3) those who perform well on LTM tasks but not AC, and 4) those who have difficulty completing both AC and LTM tasks. The presence of distinct groups would provide evidence for the idea that reading comprehension deficits may not always be due to wholesale working memory dysfunction. Indeed, a poor comprehender may be able to maintain information in focus rather well but struggle with the ability to recover this information in situations where it has been displaced from focus due to distraction or interference. Conversely, a reader may show poor comprehension performance not because they have difficulty retrieving the information that has been displaced but because they have issues maintaining that information long enough for it to be adequately processed or encoded in the first place.

CHAPTER 3
METHOD

Participants

Two-hundred and twenty-nine participants (84% female) were recruited from the research subject-pool at the University of Georgia; however, 16 participants were excluded from data analysis as they failed to complete every task. The remaining participants (*N* = 213) were between the ages of 18-37 (M = 19.65, SD = 1.84) and received course credit for their participation.

Experimental Procedure

Participants were tested either individually or with another participant in a quiet room during a single 2-hour session. All tasks were completed on a laptop computer with the exception of the reading comprehension measures, which were completed using pen and paper. Instructions for each task were provided by both a member of the research team and included on-screen before each task. After obtaining informed consent, each participant completed the WMC measures, the LTM measures, the AC measures, and the RC measures. All of the tasks were administered in the aforementioned order for all participants.

Description of Measures

Working Memory Capacity (WMC) measures

Working memory capacity was assessed using three complex span tasks: Operation span, Reading span, and Symmetry span. As discussed above, complex span tasks are used extensively in research examining individual differences in working memory abilities. In comparison with simple span tasks, complex span tasks are ostensibly dual-task measures—requiring that participants both process and store information simultaneously (for more information, see Conway, Kane, Bunting, Hambrick, Wilhelm & Engle, 2005; Unsworth, Heitz, Schrock, & Engle, 2005; Unsworth Redick, Heitz, Broadway, & Engle, 2009). Importantly, each of the complex span measures used differ in the types of stimuli presented. This is essential for reducing the possibility that working memory performance could be due to strengths (or weaknesses) in a particular domain, such as mathematics or reading. Further, shortened versions of these tasks (in which a single trial is omitted from their original versions) were employed in the current study in an effort to reduce experiment length. Recent work has indicated that this does not have a significant effect on the predictive or explanatory power of the tasks (Foster, Shipstead, Harrison, Hicks, Redick, & Engle, 2014).

1. Operation Span

Operation span (OSPAN) required participants to solve a series of math problems while also trying to remember a set of unrelated letters. Each math problem involved two required operations. The first operation always required either multiplication or division, while the second operation always required either addition or subtraction. For example, a participant may be presented with the following problem: (6 / 2) - 1 = ???. Importantly, an answer is provided directly below the problem and the participant is required to click on the True or False buttons presented on-screen, depending on whether they believe the answer provided is correct. The length of time that a participant had to answer each problem is the calculated average of their response time during an initial practice trial phase. If the participant did not make a response before time ran out or made an incorrect response, a math error was recorded. If the participant's math score fell below 85%, the task was discontinued. Feedback for providing a correct or incorrect answer was given to the participant during the practice trial phase but not during the real trial phase.

Following each math problem, participants were presented with a letter on-screen for 1 s. The letters used in each trial included: F, H, J, K, L, N, P, Q, R, S, T, and Y. Immediately after the letter was presented, the next math problem was presented. After a variable set amount of letters had been presented, participants were then asked to recall all of the letters they could remember from the current trial in correct serial order by clicking on the appropriate letters in a visual array (see Unsworth et al., 2005 for more details). The number of letters presented in each trial varied from 3 to 7, with two trials of each letter-set presented, for a total possible score of 50. Additionally, the order in which letter-sets were presented in each trial varied randomly. Participants received three trials (of letter-set size 2) of practice before beginning the real trials. A participant's recall score was the proportion of correct letters recalled in their correct serial position.

2. Reading Span

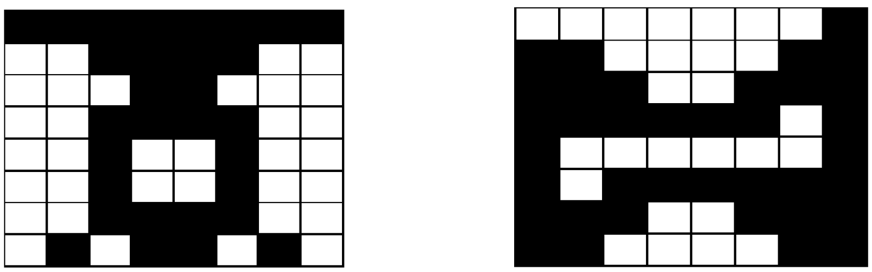
Reading span (RSPAN) required participants to read sentences while trying to remember a set of unrelated letters. For this task, the participants read a sentence and determined whether the sentence made sense by clicking True or False buttons presented on-screen. The length of time that a participant had to make their judgment was the calculated average of their response time during an initial practice trial phase. If the participant did not make a response before time ran out or made an incorrect response, a sentence error was recorded. If the participant's sentence score fell below 85%, the task was discontinued. Half of the sentences made sense, and the other half did not. Nonsense sentences were created by simply changing one word from an otherwise normal sentence (e.g., "After yelling at the game, I knew I would have a tall voice").

After participants indicated whether a sentence makes sense, they were presented with a letter for 1 s. The letters used in each trial included: F, H, J, K, L, N, P, Q, R, S, T, and Y. Immediately after the letter was presented, the next sentence was presented. After a variable set amount of letters had been presented, participants were then asked to recall all of the letters they could remember from the current trial in correct serial order by clicking on the appropriate letters in a visual array. The number of letters presented in each trial varied from 3 to 7, with two trials of each letter-set presented, for a total possible score of 50. Additionally, the order in which letter-sets were presented in each trial varied randomly. Participants received three trials (of letter-set size 2) of practice before beginning the real trials. A participant's recall score was the proportion of correct letters recalled in their correct serial position.

3. Symmetry Span

Symmetry span (SYMSPAN) required participants to recall sequences of red squares within a matrix while performing a symmetry judgment task. In the symmetry judgment task, participants were shown an 8x8 matrix with some squares filled in black (see Figure 6). Participants decide whether the design is symmetrical about its central vertical axis and indicated their response by pressing True or False buttons presented on-screen. Patterns were symmetrical half of the time. The length of time that a participant was provided to make their judgment was the calculated average of their response time during an initial practice trial phase. If the participant did not make a response before time ran out or made an incorrect response, a symmetry error was recorded. If the participant's symmetry score fell below 85%, the task was discontinued.

Immediately after determining whether a pattern was symmetrical, participants were presented with a 4x4 matrix with one of the cells filled in red for 650 ms (see Figure 7). At test, participants were asked to recall the sequence of red-square locations in the preceding trial in the serial order that they appeared, by clicking on the cells of an empty matrix. Participants received three sets (with set-size of two) of practice before beginning the real trials. Real trial sequences varied in length of 2-5. Two trials of each sequence was presented during the real trials phase, for a total possible score of 28. A participant's recall score was calculated as the proportion of correct square-locations recalled in their correct serial position.



a. *b.*

Figure 6. Example symmetry judgment items (*a*. symmetrical vs. *b*. asymmetrical) used in the symmetry span task. Participants must judge whether each image is symmetrical along its central vertical axis.

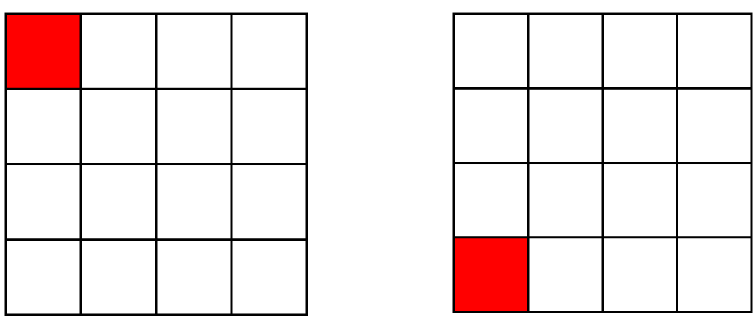


Figure 7. Example trial sequence (with set-size of 2) for memory task used in the symmetry span task. Participants must remember the location of each red-square in the order that they are presented.

Long-term memory (LTM) measures

The long-term memory measures used required participants to recall some information after a delay rather than immediately following the learning phase of that information. As such, these measures are assumed to be strong indicators of controlled retrieval abilities in the absence of active maintenance or information simply unloaded from a short-term memory store or the focus of attention. Previous work has suggested that each of these tasks measures some aspect of controlled retrieval abilities and load on the same factor (Unsworth & Spillers, 2010).

1. Delayed free recall task

During the delayed free recall tasks, participants were presented with and then asked to recall 5 lists of 10 unrelated words each. All words were common nouns, presented for 2 s each. After each list presentation, participants engaged in a 45 s distractor task before recall. The distractor task consisted of a series of three-digit numbers (e.g., 276) being presented for 2 s each and required the participant to rewrite the number beginning with the largest digit to the smallest (e.g., 762). After the distractor task, participants were asked to type as many words as they could remember from the most recently presented list in any order they wished. Participants were provided 60 s for recall before moving onto the next trial. A participant’s score was the proportion of items correctly recalled.

2. Cued recall task

The cued recall task involved presenting participants with 3 lists of unrelated words. Each list consisted of 10 pairs of words and pairs were presented on-screen for 2 s each. Participants were asked to remember the words as a pair. After each list presentation, participants were provided one word of a presented pair from the most recently presented list and asked for the associated word. Pairs were not tested in the order in which they were presented. Participants were provided 5 s to make their response before the computer moves onto the next pair. A participant's score was the proportion of pairs correctly recalled.

3. Picture-source recognition task

In the picture-source recognition tasks, pictures of objects were presented in color on-screen in one of the four corners of the computer monitor. Participants were asked to remember both the picture and the location (i.e., corner) the picture was presented in. At test, participants were presented with pictures of objects in the center of the screen and asked to identify whether the picture was "new" or "old." If old, participants were asked to press the numeric key (1, 2, 3, or 4) corresponding to the corner they believe the object was originally presented in. If new, participants were asked to press the 5 key. During the encoding phase, 30 pictures were presented at 1 s each. During the test phase, 60 pictures (30 new, 30 old) were presented and participants were provided 5 s to make their response. Additionally, order presentation of old pictures was different between encoding and test phases with new pictures randomly intermingled with old pictures during the test phase. A participant's score was the proportion of pictures correctly identified.

Attention Control (AC) measures

Attention control was assessed using the anti-saccade and Stroop tasks[1]. These tasks have been previously used to measure some aspect of attention control and do not require the

---

[1] A third task (Arrow Flankers) was originally planned to be included in the study, however, consistent program failures during an initial pilot run were experienced, resulting in a significant portion of missing data and a number of participants completing the task multiple times (leading to practice effects). Approximately 144 participants attempted the task before it was dropped from the battery (cutting the total experimental time by ~6-mins). A one-way ANOVA indicated that performance across all remaining measures between the two groups (Flankers-included vs. not-included) was not significant (*p*'s >.05), suggesting that the experimental time difference between groups did not have an overall effect on performance. While three indicators per factor is ideal for latent-variable analysis, two indicators has been found to be valid (Kline, 2016). Further, a preliminary confirmatory factor analysis indicated that Flankers did not significantly load on a latent attention control factor nor did it correlate with the other

influence of long-term memory. Previous work has suggested that each of these tasks measures some aspect of attention control and have been found to load on the same factor (Poole & Kane, 2009; Kane, Poole, Tuholski, & Engle, 2006; Unsworth, Fukuda, Awh, & Vogel, 2014; Unsworth & Spillers, 2010).

1. Anti-saccade

In the anti-saccade task (Kane et al., 2001), participants were instructed to stare at a fixation point that is on-screen for a variable amount of time (200–2200 ms). A flashing white ‘‘=” was then flashed either to the left or right of fixation for 100 ms. This was followed by a target stimulus (the letters *B*, *P*, or *R*) on-screen for 100 ms. This was followed by masking stimuli (an *H* for 50 ms and an *8* which remains on-screen until a response was given). The participants’ task was to identify the target letter by pressing a key for *B*, *P*, or *R* as quickly and accurately as possible. Importantly, response keys for *B*, *P*, or *R* were mapped (and identified with stickers) to the 1, 2, and 3 numeric keys, respectively.

In the pro-saccade condition, the flashing cue (=) and the target appeared in the same location. In the anti-saccade condition, however, the target appeared in the location opposite of the flashing cue. Participants received, in order, 10 practice trials to learn the response mapping, 15 trials of the pro-saccade condition, and 60 trials of the anti-saccade condition. The dependent variable was proportion of correct responses on the anti-saccade trials.

2. Stroop

In the Stroop task (Stroop, 1935), participants were presented with a color word (red, green, or blue) presented in one of three different font colors (red, green, or blue). The participants’ task was to indicate the font color via key press (red = 4, green = 5, blue = 6).

---

attention control tasks (*p*’s > .05). We do not interpret this finding as a valid theoretical departure but instead as the result of programming errors.

Colored stickers marked the corresponding keys. Participants were told to press the corresponding key as quickly and accurately as possible. Participants received 15 trials of response mapping practice and 6 trials of practice with the real task. Participants then received 75 total real trials. Of these trials 67% were congruent such that the word and font color matched (i.e., red printed in red) and the other 33% were incongruent (i.e., red printed in green). The dependent variable was the reaction time difference between incongruent and congruent trials.

Reading Comprehension (RC) measures

The reading comprehension measures used included the Nelson-Denny Reading Test, ACT, and the Inference Verification Task. All three of these measures have been extensively used to assess passage-level comprehension ability in adults and adolescents and require that participants read short passages and provide responses to questions about what they have read (Griffin, Wiley, & Theide, 2008; Magliano, Millis, Ozurub, & McNamara, 2007). All reading comprehension measures were implemented in paper booklet form. Further, participants were allowed to write in each booklet as they read and to make their responses, however, they were not allowed to refer back to the passage once they began answering the questions. This protocol was decided to ensure each participant carefully read each passage in full before moving onto the testing phase and that they were relying solely on their memory representation for each text when answering questions rather than simply scanning the text for keywords to make their response (a common reading comprehension test-taking strategy).

Additionally, a common practice in research on reading comprehension using latent variable analysis is to *parcel* items (or questions) from a given task to form indicators. We elected not parcel items from any given task and instead chose to use the composite scores from each of the three comprehension tests as manifest variables. The reason for this decision was two-fold. First, using multiple measures of comprehension, especially those that are commonly employed, provides a more comprehensive range of reading comprehension ability (indeed, one of the principal reasons for having multiple indicators is to reduce the task-specific variance that comes from relying upon a single indicator). Further, there is considerable debate as to whether many reading comprehension measures should be treated as though they are measuring the same skills, and so it is an interesting question theoretically whether a latent factor can be derived from a distinct array of tests (Pearson & Hamm, 2005). Second, the practice of parceling is still hotly debated, as there is some evidence that parceling may bias the model towards assumptions of unidimensionality when it may not the case (Little, Cunningham, Shahar, & Widaman, 2002). Using multiple tests as indicators, rather than parceling items within a single task, ensures that reading comprehension (and its multi-faceted nature) is adequately captured in the analysis. This is especially important if one wishes to draw conclusions about the processes involved in general constructs, such as working memory and reading comprehension.

1. Nelson-Denny Comprehension Subtest

The Nelson-Denny Reading Test is a traditional and widely employed standardized test for assessing both reading and comprehension ability. In this task, participants were asked to read five expository passages derived from the Comprehension Subtest Form-G of the Nelson-Denny Reading Test and complete five multiple choice questions (with five alternatives) following each passage. The texts varied from 157–237 words in length and had Flesch-Kincaid reading ease scores in the difficult range of 27-55. Topics included social science, humanities, history, and natural science. Participants were allowed 10 minutes to complete the entire task. A participant’s score was based on the total number of correct responses out of a possible 25.

For the purposes of the current study, adjustments were made to the normal administration procedure of the Nelson-Denny Reading Test. First, as noted above, instead of having the passage available while completing the comprehension questions, the passage was removed. Second, whereas the Nelson-Denny is typically administered without a time-limit, experimental constraints necessitated introducing one in the current study. For similar reasons, a practice passage was not given to participants, although instructions in taking the test were provided and allotted time was allotted for participants to ask for clarification. Moreover, given that all participants were college undergraduates, it was expected that they were familiar with the format of a reading comprehension test.

2. ACT Comprehension Subtest

The ACT is a popular standardized college entrance exam. The full test includes several subtests assessing science, mathematics, and reading comprehension abilities. In the current study, participants were asked to read three passages derived from the reading comprehension subtest of several practice ACT exams and answer five multiple-choice questions (with four alternatives) about each passage. The texts varied from 610–767 words in length and had Flesch-Kincaid reading ease scores in the difficult range of 44-50. Topics included social science, history, and natural science. Participants were allowed 15 minutes to complete the entire task. A participant’s score was based on the total number of correct responses out of a possible 15.

3. Inference Verification Test (IVT)

The Inference Verification Test was adapted from a similar measure developed and employed by Griffin, Wiley, and Theide (2008; see also McVay & Kane, 2012). In this task, participants were asked to read three passages and judge five statements as either “true” or “false,” based upon what was read in each text. Importantly, the statements were not paraphrases of the text but were instead assertions logically implied by the central ideas and concepts presented within each passage. Therefore, participants were required to make inferences and correctly decide whether the statement was consistent or inconsistent with the concepts presented in each text. The texts varied from 549–686 words in length and had Flesch-Kincaid reading ease scores in the difficult range of 39-58. Topics included social science, humanities, and natural science. Participants were allowed 15 minutes to complete the entire task. A participant's score was based on the total number of correct true/false judgments out of a possible 15.

CHAPTER 4

RESULTS

Descriptive statistics

Descriptive statistics are shown in Table 1. Most measures were approximately normally distributed with values of skewness and kurtosis under the generally accepted values (i.e., skewness < 2 and kurtosis < 4; see Kline, 2016). Reliability estimates were also computed using Cronbach's alpha (or coefficient alpha, α_c). This is a statistic for internal consistency, or the degree to which responses are consistent across items in a measure. Higher values indicate greater homogeneity and inter-item correlations, and is generally an indication that individual items are measuring the same content. Correlations among the measures, shown in Table 2, were weak to moderate in magnitude and, furthermore, measures of the same construct generally correlated stronger with one another than with measures of other constructs, indicating both convergent and discriminate validity within the data.

Table 1. Descriptive statistics and reliability estimates (Cronbach's alpha) for all measures.

Measure	M	SD	Skew	Kurtosis	Reliability
WMC					
Ospan	38.67	7.36	-.71	.18	.62
Rspan	37.38	7.92	-.66	.30	.69
Symspan	19.71	4.78	-.66	.66	.61
LTM					
DFR	.58	.15	.19	.43	.78
Cued	.58	.22	-.10	-.87	.79
PicSource	.76	.16	-1.26	1.45	.83
AC					
Anti	.61	.16	.09	-.78	.81
Stroop	120.12	71.55	.62	.25	NA
RC					
ND	.79	.10	-.57	-.05	.56
ACT	.68	.11	-.56	.43	.44
IVT	.77	.11	-.57	.62	.34

Note: Latent Factors: WMC = working memory capacity; LTM = long-term memory retrieval; AC = attention control; RC = reading comprehension. Indicators: Ospan = operation span; Rspan = reading span; Symspan = symmetry span; DFR = delayed free recall; Cued = cued recall; PicSource = picture-source recognition; Anti = antisaccade; Stroop = Stroop; ND = Nelson-Denny; ACT = ACT; IVT = inference verification task

Table 2. Correlations between all measures.

Measure	1	2	3	4	5	6	7	8	9	10	11
WMC											
1. Ospan	-										
2. Rspan	.52*	-									
3. Symspan	.42*	.44*	-								
LTM											
4. DFR	.24*	.31*	.27*	-							
5. Cued	.13	.22*	.17*	.47*	-						
6. PicSource	.17*	.28*	.42*	.34*	.35*	-					
AC											
7. Anti	.14*	.13	.23*	.15*	.18*	.21*	-				
8. Stroop	-.11	-.14*	-.15*	-.10	-.06	-.05	-.18*	-			
RC											
9. ND	.12	.29*	.13	.25*	.41*	.30*	.07	-.01	-		
10. ACT	.06	.24*	.17*	.26*	.26*	.21*	.08	-.05	.44*	-	
11. IVT	.10	.16*	.06	.18*	.18*	.11	.02	-.10	.30*	.34*	-

Note: Latent Factors: WMC = working memory capacity; LTM = long-term memory retrieval; AC = attention control; RC = reading comprehension. Indicators: Ospan = operation span; Rspan = reading span; Symspan = symmetry span; DFR = delayed free recall; Cued = cued recall; PicSource = picture-source recognition; Anti = antisaccade; Stroop = Stroop; ND = Nelson-Denny; ACT = ACT; IVT = inference verification task. *Correlations >.14 are significant at the $p < .05$ level.

Confirmatory factor analysis

First, confirmatory factor analysis was used to test several measurement models to determine the structure of the data. As discussed previously, confirmatory factor analysis provides a means of testing the adequacy of competing theoretical accounts of the relationships between measures, with each model specified in terms of paths between indicators (also called observed variables) and latent constructs, and between constructs themselves (Bollen, 1989; Kline, 2016). Specifically, four initial measurement models were specified to determine how WMC, AC, and LTM were related to one another. These models were tested to verify that the selected indicators had significant path loadings on their hypothesized latent factors and, further, to replicate previous research on the dual-component model suggesting WMC, AC, and LTM are correlated, yet distinct factors. A fifth measurement model was then tested to include a latent factor for RC ability. This model was tested to provide a basis for our proposed structural model and to begin answering the primary question of interest—where do the AC and LTM factors figure into the relationship between WMC and RC?

Each measurement model was scaled according to unit loading identification (ULI), in which the factor loading of the first indicator for each latent factor was fixed to 1.0 (otherwise known as the reference variable). This is in contrast to the scaling method of unit variance identification (UVI), wherein the variance of each latent factor is fixed to 1.0. Scaling methods are simply tools for achieving model identification. Further, the decision to utilize the ULI scaling method over UVI was mostly arbitrary as it is the default scaling method for *R* (and most other statistical software packages used to conduct SEM analysis). Further, in studies where the CFA is analyzed in a single sample and there are no repeated measures variables, as is true in the current study, either method is appropriate (Kline, 2016). Additionally, standardized path coefficients were used for reporting factor loadings, as depicted in Figure 8. This was done to allow for ease of comparison of the factor loadings for each indicator and to stay consistent with previous research involving latent variable analysis and individual differences in working memory capacity (e.g., Unsworth & Spillers, 2010). Note standardized path coefficients does not equate to amount of variance accounted for.

Model 1 tested the notion that WMC, AC, and LTM abilities are best conceptualized as a single unitary construct, consistent with the idea that there is a unitary controlled attention or central executive factor responsible for completion across all tasks (Engle et al., 1999; Alloway, Gathercole, & Pickering, 2006; Alloway, Gathercole, Kirkwood, & Elliot, 2009). Thus, in this model, all of the attention and memory tasks were loaded onto a single factor. Model 2 tested the idea that WMC and AC are largely the same construct but distinct from the LTM factor, consistent with the notion that working memory performance largely reflects attention abilities distinct from basic memory abilities. Model 3 tested the notion that WMC and LTM reflect the same abilities and thus load onto a single factor separate from the AC factor (Mogle et al., 2008). Model 4 tested the notion that the three constructs, WMC, AC, and LTM are best thought of as three related yet distinct factors (Unsworth & Spillers, 2010; Unsworth et al., 2014). As discussed above, recent research from our lab has found the both AC and LTM account for unique variance in WMC, therefore, we expected Model 4 as having the best fit. Finally, a fifth model (Model 4-RC) was tested to observe whether the comprehension measures would load on a separate latent RC factor and if there were correlations between this new variable and the other three factors. We anticipated the WMC factor to be related to the RC factor as well as the AC and LTM factors to be related to the RC factor, provided the dual-component model predictions hold true. Once established, Model 4-AC was planned for use in the later structural analysis.

Model fits were assessed using a combination of several fit statistics. These included chi-square (χ^2), root mean square error of approximation (RSMEA), standardized root mean square residual (SRMR), the non-normed fit index (NNFI), the comparative fit index (CFI), and the Akaike information criterion (AIC). The chi-square statistic is considered an *exact-fit* statistic and reflects whether there is a significant difference between the observed and reproduced covariance matrices. Therefore, non-significant values are desirable. The RMSEA is an index of model misfit due to model misspecification and the SRMR reflects the average squared deviation between the observed and reproduced covariances. In addition, the NNFI and CFI compare the fit of the specified model to a baseline null model. RMSEA and SRMR values < .08 and NNFI and CFI values > .90 are indicative of acceptable fit (Kline, 2016). Finally, the AIC examines the relative fit between models in which the model with the smallest AIC is preferred. It is important to examine multiple indices because ultimately model comparisons are an approximation, rather than an exact science. Each of the indices mentioned above reflect only one aspect of fit and are, therefore, insufficient in isolation for making decisions about whether to retain a model. Moreover, fit statistics can aid the decision to accept, reject, or modify a model but they cannot provide strict yes-or-no answers. Instead, it is important to consider the fit statistics within the context of theory; that is, how does my hypothesized model compare to reasonable alternatives? Including multiple fit statistics also precludes one from cherry-picking one over another for the purposes of advancing one model over another. Finally, the described threshold values for what is considered “acceptable” fit are simply rules-of-thumb derived from computer simulation studies. They were not designed to act as strict cutoffs or used to confirm the legitimacy of the hypothesized model.

Shown in Table 3 is the fit of the five tested measurement models. As can be seen, Measurement Model 4 that specified three separate, yet correlated, factors (WMC, AC, and LTM) provided the best fit, replicating previous research arguing that both AC and LTM abilities account for variance in WMC. Specifically, Model 4 fit significantly better than the three alternative models (all $\Delta\ \chi^2$’s > 9.5, p’s < .01). Importantly, the fit of Model 4-RC was also good, and had the lowest AIC value, providing justification for the structural model analysis. As seen in Figure 8, nearly all of the indicators loaded significantly onto their respective factor and most all of the factors were moderately interrelated. These results suggest that WMC, LTM, and AC are best thought of as separate, yet correlated factors, each of which is related to RC.

Notably, the Stroop task loaded weakly onto the AC factor and, further, the AC factor was not significantly correlated to the RC factor.

Table 3. Fit indices for all tested confirmatory factor analysis and structural equation models.

Model	χ^2	*df*	RMSEA	SRMR	NNFI	CFI	AIC
Model 1	74.64	20	.11	.07	.72	.80	5866.99
Model 2	38.52	19	.06	.05	.89	.93	5832.87
Model 3	71.22	19	.11	.07	.72	.81	5865.57
Model 4	32.89	17	.06	.05	.90	.94	5831.24
Model 4-RC	56.98	38	.04	.04	.93	.95	4780.42
SEM	59.29	39	.04	.05	.92	.94	4780.73

Note. RMSEA = root mean square error of approximation; SRMR = Standardized Root Mean Residual; NNFI = non-normed fit index; CFI = comparative fit index; AIC = Akaike information criterion.

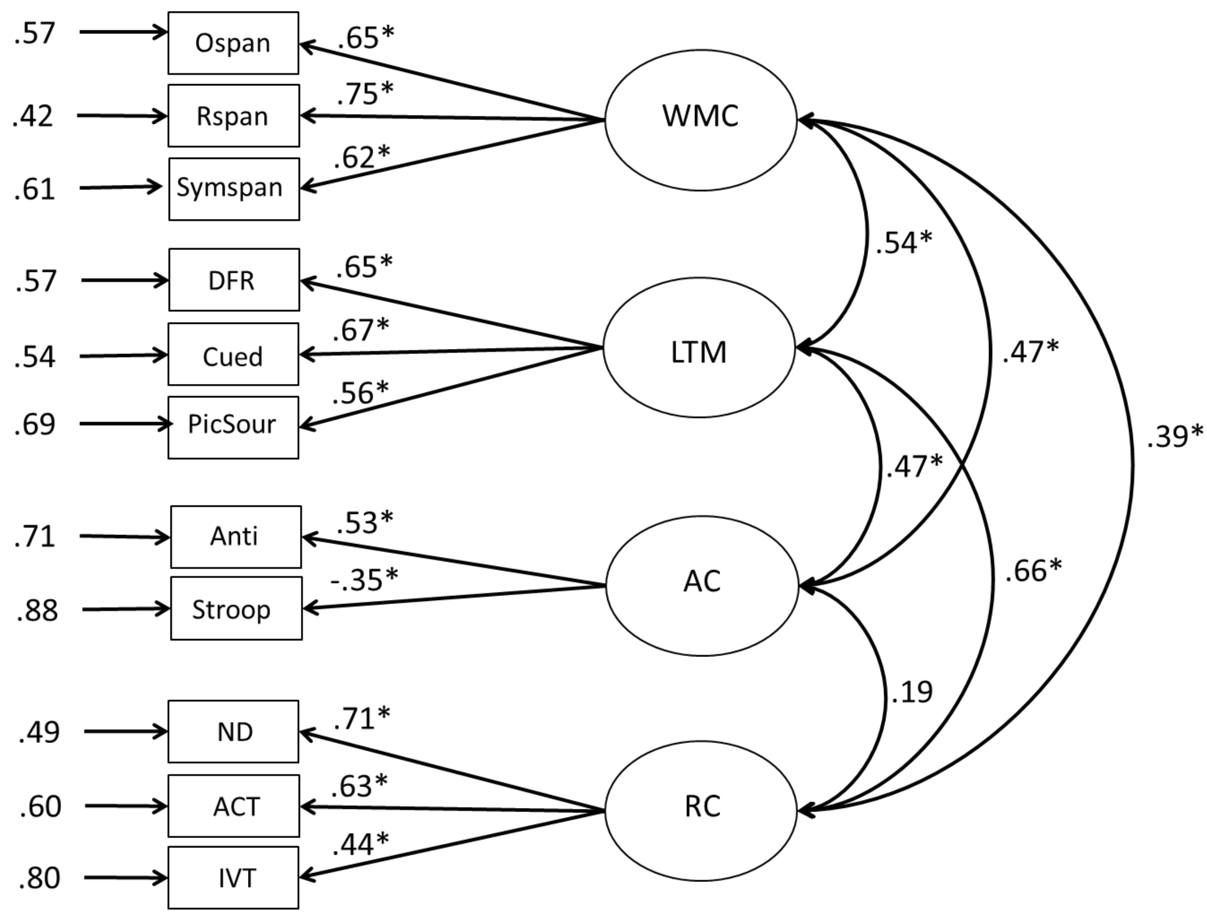


Figure 8: Graphical depiction of the fifth tested measure model, Model 4-RC. Paths connecting latent factors (circles) to each other represent correlations between factors, the numbers from the latent factors to the indicators (squares) represent the factor loadings of each task onto the latent factor, and numbers appearing next to each indicator represents error variance associated with each task. Latent factors: working memory capacity (WMC), long-term memory (LTM), attention control (AC), and reading comprehension (RC). Indicators: Ospan = operation span; Rspan = reading span; Symspan = symmetry span; DFR = delayed free recall; Cued = cued recall; PicSour = picture-source recognition; Anti = antisaccade; Stroop = Stroop; ND = Nelson-Denny; ACT = ACT; IVT = inference verification task. * indicates significant correlation at the $p < .05$ level.

Structural equation model analysis

The results of the CFA suggested WMC, LTM, AC, and RC could be considered as four distinct factors, therefore, we utilized structural equation modeling to examine the primary question of interest—namely, which specific components of WMC underlie its relation to RC? Specifically, a model was tested (see Figure 9) to see if there were indirect effects of the LTM and AC factors in the relationship between WMC and RC. Indirect effects are best described as effects in which an initial variable (X_1) causes some change in another, intervening variable (Y_1) which in turn leads to changes in the outcome variable (Y_2). The variable X_1 causes effects in Y_2 indirectly through Y_1. Importantly, indirect effects should be distinguished from true mediation. In fact, mediation is a subclass of indirect effects that involves observing actual *changes* in variables. In cross-sectional designs, however, where data is collected all at once rather than in a longitudinal fashion, it can be difficult to infer mediation effects between causal and outcome variables. Therefore, interpretation of the results as indirect effects, rather than as mediation, is more appropriate.

If the predictions of the dual-component model are correct and the LTM and AC factors do indirectly affect all (or most) of the relation between WMC and RC, we should see that (1) WMC predicts each intervening factor, (2) that each intervening factor significantly predicts RC, and (3) that WMC no longer has a significant direct effect on RC. As shown in Table 3, the fit of the resulting model was good. Further, the results of this analysis indicated that WMC significantly predicted both the LTM and AC factors suggesting that WMC is uniquely related to each of these factors (Unsworth & Engle, 2007; Unsworth & Spillers, 2010). Notably, however, only the LTM factor significantly predicted RC suggesting that only LTM, and not AC, contributes to variation in RC. Most importantly, the direct path from WMC to RC was not significant. Indeed, the correlation between WMC and RC went from $r = .39$ to approximately zero after statistically controlling for other factors. Thus, the relationship between LTM and WMC was fully explained through the indirect effect of LTM, as once this factor was taken into account WMC no longer predicted residual variance in RC.

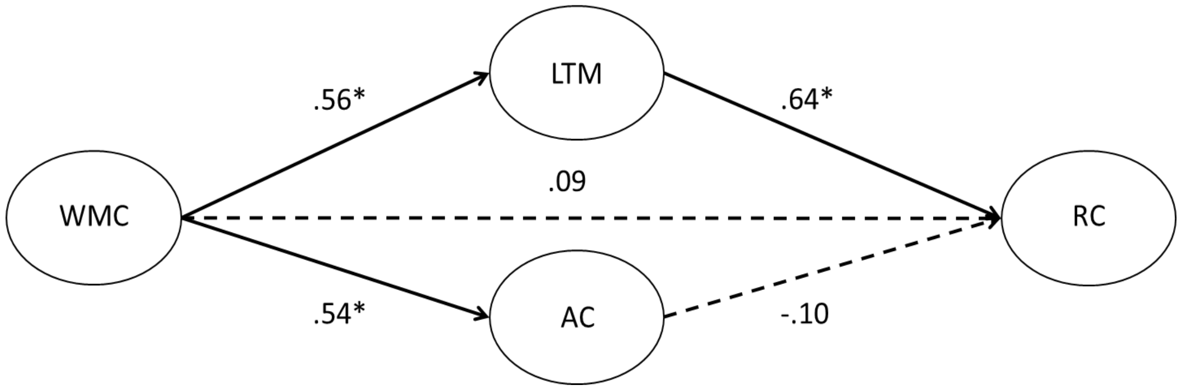


Figure 9. Structural equation model for working memory capacity (WMC), long-term memory (LTM), attention control (AC), and reading comprehension (RC). Single-headed arrows connecting latent variables (circles) to each other represent standardized path coefficients indicating the unique contribution of the latent variable. Solid lines are significant at the $p < .05$ level while dotted lines are not significant at the $p <.05$ level. Note: Indicator variables have been omitted for clarity.

Sub-group analysis

The final set of analyses conducted was to categorize participants into sub-groups according to specific patterns of performance across the long-term memory and attention control measures. The purpose for this analysis was to gain a more nuanced perspective as to why exactly individuals with working memory deficits struggle with comprehending what they are reading. Indeed, if working memory is multi-faceted, it is an important question to ask whether comprehension issues are due mostly to weaknesses in attention control or long-term memory retrieval abilities, so that more tailored interventions could be developed.

To conduct the sub-group analysis, factor composites for the LTM and AC measures were formed and entered into a two-step cluster analysis using SPSS. In this analysis, algorithms group cases into pre-clusters by constructing a cluster feature tree and then determine whether a case should be included with a previously formed pre-cluster or whether a new pre-cluster should be created based on the cluster feature tree. The second step involves using a hierarchical clustering method to determine the number of distinct clusters present in the data. In this stage, clusters are recursively merged until the desired number of clusters is determined. The distance between clusters is based on a log-likelihood measure, whereby distance is related to the decrease in log likelihood as the clusters are formed into a single group. The algorithm automatically determines the number of distinct groups by taking into account the lowest information criterion (AIC) and the highest ratio of distance measures (indicating the best separation of the groups). Once groups were determined, factor composites were computed for the WMC and RC factors, so that performance comparisons between groups could be made on these measures as well.

Recall that if both AC and LTM can independently affect performance on measures of reading comprehension, then it is possible that readers can fall into distinct groups based on their respective AC and LTM abilities. Specifically, we predicted four different patterns of performance: 1) those who perform well on both AC and LTM tasks, 2) those who perform well on AC tasks but not LTM, 3) those who perform well on LTM tasks but not AC, and 4) those who have difficulty completing both AC and LTM tasks. Indeed, the sub-group analysis revealed the presence of four distinct groups in the data—consisting of 40, 56, 27, and 90 participants, respectively. Shown in Table 4 are the mean composite scores for each group across each factor.

At first glance, there are discernable patterns of performance differences between groups across each factor. While the results of an omnibus one-way ANOVA confirmed these initial observations (see Table 4), Bonferroni post-hoc comparison tests were conducted in order to examine the specific ways the groups differed with respect to each factor. Examining LTM performance, it was found that Groups 1 and 3 were strong in long-term memory retrieval abilities (and not significantly different from each other, $p > .90$), while Groups 2 and 4 were significantly weaker (all p's $< .05$). A difference was also observed in performance between Groups 2 and 4, such that Group 2 performed significantly worse than Group 4 ($p < .05$). With respect to AC performance, it was found that there were significant differences between all groups—but with Group 1 performing the worst overall and Group 2 performing the best overall when compared to all other groups (all p's $< .05$).

Importantly, the results of these comparison tests were consistent with our initial hypotheses—namely, that it is possible for individuals to show differential deficits or strengths in the components of working memory, rather than an overall working memory deficit or strength. Specifically, Group 1 consisted of participants that demonstrated high LTM ability but lower than average AC ability, while Group 2 consisted of participants with the weakest LTM scores, yet the strongest AC scores. Group 3 consisted of relatively good performers overall, showing both strong LTM scores and above average AC scores. Conversely, Group 4 (the largest sub-group) consisted of participants that could be considered fairly average in performance across all factors.

Once groups had been defined and it was apparent that they showed differences in performance in each of the components of working memory, the next step was to determine if these groups also differed on measures of WMC and RC. That is, could it be possible that a certain component of working memory (i.e., long-term memory retrieval vs. attention control) be more advantageous for WMC and RC performance? Or are both equally beneficial? After conducting post-hoc comparison tests, it was found there were significant differences between the groups in WMC performance. Specifically, Groups 1 and 3 performed well above average and not significantly different from each other ($p > .90$). Group 4 performed average and was significantly different from Groups 1 and 2 (both p's $< .05$), but not significantly different from

Group 3 ($p > .70$). Group 2 performed rather poorly compared with all other groups (all *p*'s < .05). Additionally, performance differences were observed in RC across the groups as well, though the differences were not as stark. Indeed, Groups 1 and 3 did not significantly differ from each other (though Group 3 performed best overall) but did significantly differ from Groups 2 and 4. Group 2 performed significantly poorly compared with the other groups (with the exception of Group 4, where no significant difference was observed, $p > .80$).

Taken together, these results indicate that reading comprehension deficits could be due to weaknesses in either component of working memory, rather than wholesale dysfunction. The results also suggest that an inability to retrieve information from long-term memory will have a greater negative effect on an individual's reading comprehension performance than a comparable weakness in the ability to control their attention in the face of distraction while reading. This was apparent from the lack of significant differences in RC performance between Groups 1 and 3 (groups that performed well overall on LTM measures), whereas Group 2 showed significant RC deficits despite having the strongest performance on AC measures. Moreover, having both strong LTM *and* AC abilities appears to be more advantageous than being skilled in either alone, as suggested by the above average overall performance of Group 3 compared with all other groups.

Table 4. Mean factor composite scores and sub-group analysis results.

Measure	Group 1	Group 2	Group 3	Group 4	*F*	η^2
N	40	56	27	90	—	—
LTM	.91 (.06)	- .63 (.08)	.97 (.09)	- .30 (.04)	114.4	.67
AC	- .50 (.04)	.65 (.04)	.31 (.04)	- .27 (.03)	144.9	.62
WMC	.45 (.09)	- .44 (.13)	.26 (.15)	- .01 (.07)	10.8	.13
RC	.31 (.12)	- .27 (.11)	.36 (.10)	- .07 (.07)	6.7	.09

Note. *N*= size of group membership; LTM = factor composite of three long-term memory retrieval measures; AC = factor composite of two attention control measures; WMC = factor composite of three working memory capacity measures; RC = factor composite of three reading comprehension measures. Numbers in parentheses are standard errors of the mean. All *F*'s significant at $p < .01$.

CHAPTER 5

DISCUSSION

The primary purpose of the current study was to investigate the relationship between working memory and reading comprehension and the specific control processes that underlie this relationship using latent variable analysis. Specifically, we sought to examine the relative influence of each component of the dual-component model of working memory on the relationship between working memory and reading comprehension. That is, is the relation primarily driven by the ability to maintain focus on task-relevant information (attention control), retrieve information from long-term memory that has been lost from focus or is task-related (long-term memory retrieval), or are both just as important?

Latent variables of working memory capacity, long-term memory retrieval, attention control, and reading comprehension performance were constructed based on multiple measures of each construct. Using confirmatory factor analysis and structural equation modeling, we replicated previous findings that working memory capacity is correlated with reading comprehension ability and that both attention control and long-term memory retrieval are significantly related to working memory capacity. Importantly, we also found that it was long-term memory retrieval, and not attention control, that predicted variance in reading comprehension performance. In fact, the results of the SEM analysis indicated that there was an indirect effect of long-term memory retrieval in the relation between working memory capacity and reading comprehension such that the direct effect was no longer significant. A comparable indirect effect between attention control and working memory capacity was not observed.

Finally, a sub-group analysis revealed that not all participants could be classified in terms of overall deficits. Rather, sub-groups of participants were found within the data, with each group exhibiting specific deficits or strengths in the different components of working memory. Specifically, one group could be characterized as high performing overall, with strong scores on both the long-term memory retrieval and attention control measures. Conversely, another group was decidedly average across both measures. Another group showed a high ability for long-term memory retrieval, but weaknesses in tasks requiring attention control. Finally, the last group showed a high ability for attention control, but weaknesses in tasks requiring long-term memory retrieval.

Importantly, these four groups also tended to differ in working memory and reading comprehension performance, with some groups having lower working memory capacity and comprehension scores than other groups. This suggests that there are likely multiple sub-groups of low-ability individuals, and that a sweeping generalization as to why a student may be underperforming could be problematic. Indeed, some individuals may have issues with maintaining focus on the current task or handling multiple pieces of information in mind at once, while others may show difficulty remembering what has just been read and connecting new information with prior knowledge. That said, the results also appear to demonstrate an individual's ability (or inability) to process information in the short-term is ultimately a non-issue as long as they are adept at retrieving any information that has either fallen out of focus or is important for constructing a coherent mental representation of what they are reading (i.e., generate a situation model).

On the importance of retrieval processes for memory and comprehension

Collectively, the results suggest two conclusions concerning the working memory processes that are engaged during the act of reading comprehension: 1) that an individual's attentional "capacity" may be a non-factor for comprehension performance and 2) that comprehension is driven by readers using retrieval processes to access task-relevant information stored in long-term memory in an effort to establish a situation model. Though somewhat unexpected, these results are not altogether surprising when one considers the limited nature of the focus of attention in working memory. Recall Cowan (2001)'s embedded process model of working memory, which posits that individuals can actively maintain only 3-5 pieces or "chunks" of information at a time. In order to support complex cognition (such as reading comprehension), the embedded process model argues that the bulk of task-relevant information and processing is off-loaded outside the focus of attention and that only a portion of what is being read is being maintained in focus at any given time as a reader moves through the text. Similarly, Ericsson and Kintsch (1995)'s long-term working memory model proposes that text comprehension relies almost entirely on long-term memory retrieval processes and that "the transient portion of working memory [i.e., short-term maintenance] is not necessary for continued comprehension" (p. 224-225). Only when new elements of a text are encountered and must be integrated into the existing representation already stored in long-term memory, do the authors argue concede short-term storage and attentional processing is necessary—and even then, the amount of information being attended to is still minimal due to capacity constraints (see also Delaney & Ericsson, 2016).

Moreover, the results suggest that retrieval processes must be in play *during* the learning/encoding phase (i.e., while the individual is still reading through the text)—not just at test. Indeed, given the limited nature of the focus of attention, retrieval appears necessary early on in order for the reader to establish connections within the text, chunk gist information, create retrieval cues, and ultimately generate a situation model. In the event that a reader is interrupted or distracted (i.e., the focus of attention is directed away from task-relevant material), the comprehension process will take some time to resume, as the content of new sentences will be processed without the benefit of existing context and background information. In other words, a reader may need to re-read sentences to determine where they left off and regain their train of thought (i.e., retrieve the task-goal and pertinent content information back into focus). However, once the reader has reinstated contextual information and cues from long-term memory, the comprehension process should proceed without issue and at the baseline pace (determined by both the individual's reading ability and overall text complexity). Later on, when the reader is tested for understanding of what they have read, retrieval processes will again be necessary for success.

One may wonder how it is possible that retrieval processes can have greater influence on comprehension performance beyond the encoding of information. Indeed, the present results seem counter-intuitive since the material that is being read and comprehended must first be encoded before it can be retrieved—and thus, leading to the initial assumption that attention control must be the more integral component of working memory during comprehension. Recent research examining *retrieval practice effects*, however, lends significant insight into why retrieval has such a profound effect on an individual's ability to learn and comprehend what they are reading. Indeed, the retrieval practice effect (also known as retrieval-based learning) refers to the general finding that repeated retrieval of to-be-remembered material enhances understanding and memory for said material beyond repeated studying (Karpicke, Lehman, &

Aue, 2014; Roediger & Karpicke, 2006a, 2006b; Zaromb & Roediger, 2010). This an important phenomenon because retrieval processes have historically been considered merely as an assessment of what has been learned rather than thought to actually *produce* learning themselves (Roediger & Nestojko, 2015). In standard retrieval practice paradigms, a set of to-be-remembered material is encoded and then subsequently retrieved. After the initial testing period participants engage in some other activity or delay before having their memory for that same information probed again later in the future. Memory for intermediately retrieved information has been found to be less prone to forgetting and more accessible for future retrieval attempts, compared with conditions wherein participants are simply re-presented the material or allotted additional opportunities to re-study (Karpicke & Roediger, 2007). In other words, the act of retrieving information from long-term memory is more beneficial for performance and learning than simply maintaining attention to that information over the same duration.

When applied to reading comprehension, actively retrieving information during the encoding of text benefits the reader in two specific ways. First, each successful recovery attempt of information from long-term memory increases the strength between a given retrieval cue and a memory trace as well as the diagnostic value of these cues (Karpicke, Lehman, & Aue, 2014; see Raaijmakers & Shiffrin, 1992; Nairne, 2002 for a discussion of cue-dependent search processes). Second, retrieval facilitates creation of the situation model that is necessary for true comprehension by allowing the reader to connect ideas and concepts derived from earlier portions of the text with new sections. In the event that the reader fails to retrieve pertinent information as they read through the text, overall text coherence will suffer, as will performance on later comprehension assessments. Indeed, the sub-group analysis showed that individuals with high ability in long-term memory retrieval and low ability in attention control performed significantly better on measures of reading comprehension than compared with the inverse group. In fact, it could be that the act of losing focus and having to retrieve information back into mind is helping this group compensate for their lack of ability to control attention. Essentially, these individuals undergo significant retrieval practice simply by virtue of their cognitive weakness—an inability to maintain focus on task-relevant information. On the other hand, if a reader is merely passively processing information in the short-term but not retrieving the information from long-term memory in an effort to establish connections between propositions in the text, comprehension will suffer such that on a later test, the connections between information will not be as “practiced” or accessible. In other words, the reader will have failed to generate an adequate situation model and instead must rely on the less comprehensive text-base or even simply the surface model to answer questions.

Recommendations for educators

What do the current results mean for the classroom? First, going forward it is important that educators are careful to attend to which stage the comprehension process begins to break down for a student. Indeed, an important finding in the current study was that individuals could potentially fail to comprehend what they are reading because of weaknesses in one component of working memory rather than a general deficit. Therefore, when presented with a student that is exhibiting comprehension issues, it is important to ask whether the information too complex for the student to manage, such that the student has to continuously re-read? Or whether the student is failing to retrieve (i.e., forgetting) integral pieces of information when it comes time to integrate what has been read or answer questions on a test?

Further, considering the pivotal influence that retrieval abilities have on comprehension performance, it is recommended that educators develop activities and opportunities that encourage students to actively retrieve to-be-learned material, rather than simply providing additional study exposure. Teachers could emphasize the need to recall earlier information as students are reading texts and encourage them to retrieve related prior knowledge as they guide students in the process of building up a representation of what they are reading. When using informational texts, they can provide related texts that require students to integrate information from a currently read text with knowledge of texts they have recently read. Further, teachers can create teaching contexts that encourage retrieval of information they have learned from texts by frequent quizzing or questioning about text content. For example, a recent study by McDaniel, Thomas, Agarwal, and Roediger (2013) explored the effects of frequent quizzing of science concepts and found that students performed better on topics that were quizzed than on non-quizzed topics. Similarly, Karpicke and Blunt (2011; see also Blunt & Karpicke, 2013**)** found that students who practiced retrieval of educational material performed better on later assessments than those that simply read the material repeatedly or created elaborate concept maps.

Limitations of the current study

Although we feel confident in the finding that long-term memory retrieval is integral to comprehension performance, we are reluctant to conclude that attention control plays no role in the relationship between working memory and reading comprehension. Indeed, the current results stand in stark contrast to previous studies supporting the theory that attention control and active maintenance abilities are necessary for successful comprehension (cf. McVay & Kane, 2012; Unsworth & McMillan, 2013). Further, one could imagine situations in which the ability to maintain focus in the face of distraction would prove beneficial for successful comprehension, such as experiencing multiple or prolonged disruptions while reading. In fact, there is a growing body of research showing that an interruption during reading significantly disrupts the comprehension process such that performance on later tests declines. Specifically, Foroughi and colleagues (Foroughi, Werner, Barragán & Boehm-Davis, 2015; Foroughi, Barragán & Boehm-Davis, 2016; although see Delaney & Ericsson, 2016 for a rebuttal of these findings) have found that consistent interruptions following reading a passage has a negative effect on comprehension performance, when later tested with a multiple-choice test that requires readers to recall information rather than simply rely on recognition skills. The authors argue interruptions essentially disrupt the short-term processing component of working memory, which leads to poorly encoded representations in long-term memory (affecting *accessibility* of the information for retrieval) or even a complete failure to encode that information at all (affecting the *availability* of the information for retrieval). As mentioned previously, both the embedded process model of working memory and the C-I model assign associative functions to the focus of attention, such that it is essential for the synthesis and integration of information and the creation of retrieval cues. Therefore, any disruption in this binding or associative process could potentially affect the integrity of the situation model, such that it becomes incomplete or less coherent. Readers that can simultaneously maintain multiple complex propositions in mind will likely generate a richer situation model for what they have read compared with those who struggle with capacity-overload. The current finding that the group of participants that scored high on both the long-term memory retrieval and attention control measures consistently outperformed those with strengths in only one or the other component supports this notion. Additional research will be necessary to further test these ideas.

We believe the lack of an observed effect for attention control in our latent variable analysis may also be due to poor specification of the latent AC factor. Indeed, the intercorrelations between the tasks chosen to measure attention control were rather weak and, due to the aforementioned task failures, our AC factor consisted of only two indicators, rather than the preferred three. One reason for the poor correlation between the attention control measures may be that we have treated our AC factor as a unitary attention construct, when, in fact, attention has been argued to consist of a variety of distinct processes and behaviors (Woodman, Vogel, & Luck, 2001; Cowan; Fristoe, Elliot, Brunner, & Saults, 2006). While the expectation when utilizing latent variable analysis is to eliminate task-specific variance and focus only on the "overlap" among measures, it is possible that our attention control tasks were too inherently disparate in terms of the processes or abilities they purport to measure. For instance, some tasks may measure attention capacity in terms of storage or the amount of information that can be maintained, while other attention tasks are actually measuring the capacity to do something such as keep information in an active state (or in "focus) while preventing or inhibiting irrelevant and distracting stimuli from interfering. Still other tasks may measure the ability of a participant to switch the focus of their attention from one stimulus to another or require attention to be sustained over a given period of time. Our theoretical conception and discussion of attention control in the current paper has encompassed both *capacity* (e.g., the bottleneck hypothesis and the embedded-process model) and *control* (e.g., controlled attention account of working memory capacity) views of the construct, but the tasks employed may have been more representative of the latter. Going forward, researchers interested in measuring the influence of attention and working memory on reading comprehension performance should be more selective of the specific tasks that are employed to estimate the AC factor and the actions they require for success. Indeed, an interesting question would be whether AC should be differentiated into two factors—one measuring more "capacity" aspects of attention and another that measures "control"—and whether these two factors differentially affect the comprehension process.

Correlations among the RC tasks were also lower than expected. This is surprising given that prior research has reported a much stronger relation between the ND and IVT tasks (r = .64; Griffin, Wiley & Theide, 2008). At present, the discrepancy between the current data set and those published previously is unclear. However, like the attention control tasks, there is reason to believe that these two tasks could be measuring separate things. The Nelson-Denny has recently come under fire for the nature of its questions. Specifically, researchers have argued that the Nelson-Denny tests primarily surface level facts that may be influenced more by memory and prior knowledge than actual comprehension (Wiley, Griffin, & Thiede, 2005; Griffin, Wiley, & Thiede, 2008; Magliano, Millis, Ozuru, & McNamara, 2007; Kintsch, 1998). The IVT, on the other hand, was specifically designed to avoid this confound by testing "deeper" concepts that would require a reader to reason to answer the question rather than simply retrieve or look up that information. This could explain why the Nelson-Denny generally had stronger correlations with the LTM tasks compared with the IVT. Though the correlations were low, each task significantly loaded on the latent RC factor and, further, we believe it is important theoretically that these tasks are included in the proposed study given their ubiquitous nature, generally. The ND and ACT are both heavily used in both research and classroom settings to assess reading and comprehension abilities in adolescents and adults. If anything, the lack of correlation (not to mention the less than optimal reliability) speaks to the caution that researchers should exhibit when using these tasks in the first place.

Finally, the lack of correlation between the attention control and reading comprehension factors could be the result of our comprehension measures not being sensitive enough to the stage of comprehension in which attention control is important. Indeed, we have alluded to the idea that attention control is mostly important early on in the comprehension process as the reader begins collecting and associating pieces of information and propositions from the text. However, once initial connections have been established (i.e., the text base has been formed), we have argued that it is primarily retrieval processes that assist in generating the situation model. Due to the nature of our comprehension tests, they ostensibly take place after the comprehension process has been presumed to occur—thus, leaving retrieval of this information as the primary mechanism for successfully answering the test questions. While we feel the results supporting the influential role of long-term memory retrieval in the comprehension process are valid, future studies should aim to find better methods for indexing the recruitment of attention control processes during the comprehension process (for instance, using TUTs). It is also important that future research consider the effects of having the text available to participants during test compared with making it unavailable (as we have done in the current study). Indeed, it is possible that having the test questions in mind while reviewing the text would help participants focus the connections they create between content. When readers have a particular goal or question in mind, it may alter the strategy with which they process the information they are reading. Researchers must be careful, however, to avoid the potential confound, wherein participants simply scan passage content for key words and ideas relevant to the test questions, without first attempting to read and synthesize the entire passage.

Conclusion

The current study is novel in its attempt to delineate the component processes of working memory involved in successful text comprehension. Using both a latent variable and sub-group analysis revealed two main findings: 1) long-term memory retrieval influences reading comprehension performance beyond that of attention control and 2) failures to comprehend what is being read could be the result of weaknesses in specific components of working memory, rather than an overall working memory deficit. We have argued that active retrieval of information from long-term memory promotes successful comprehension and the generation of the situation model by strengthening the connections between concepts and establishing the diagnostic value of retrieval cues. Therefore, researchers and educators interested in enhancing reading comprehension among students should focus on ways to promote retrieval of to-be-learned material during the reading of complex text materials, rather than simply providing the opportunity to re-read. Finally, while these results are important for our understanding of reading comprehension, it was nevertheless unexpected to find that attention control played little to no role in predicting reading comprehension performance, especially since this is in contrast to previous studies. Further research is necessary for determination of this issue.

REFERENCES

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131*, 30-60.

Alloway, T. P. (2006). How does working memory work in the classroom? *Educational Research and Reviews, 1(4),* 134-139.

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child development*, *80*(2), 606-621.

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuo-spatial short-term and working memory in children: Are they separable? *Child Development, 77*, 1698–1716.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261-295.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In Kenneth W. Spence and Janet Taylor Spence (Eds.). Psychology of Learning and Motivation (Vol. 2, pp. 89-195). New York: Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American, 225*, 82–90.

Baddeley, A. D. (1979). Working memory and reading. In *Processing of Visible Language* (pp. 355-370). Springer: US.

Baddeley, A. (1981). The concept of working memory: A view of its current state and probable future development. *Cognition*, *10*, 17-23.

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

Baddeley, A. (1992). Working memory. *Science*, *255*(5044*)*, 556.

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

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.). The Psychology of Learning and Motivation (Vol. 8, pp. 47–89). New York: Academic Press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: the multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory* (pp. 28–61). Cambridge, U.K.: Cambridge University Press.

Baddeley, A., Logie, R., Nimmo-Smith, I., & Brereton, N. (1985). Components of fluent reading. *Journal of memory and language*, *24*(1), 119-131.

Bandalos, D.L. & Boehm-Kaufman, M.R. (2008). "Four common misconceptions in exploratory factor analysis". In Lance, Charles E.; Vandenberg, Robert J. Statistical and Methodological Myths and Urban Legends: Doctrine, Verity and Fable in the Organizational and Social Sciences. Taylor & Francis: pp. 61–87.

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail working memory and “choking under pressure” in math. *Psychological Science*, *16*(2), 101-105.

Blunt, J. R., & Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology*, *106*(3), 849.

Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons.

Broadbent, D.E. (1958) *Perception & communication.* New York: Pergamon Press.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental neuropsychology*, *19*(3), 273-293.

Cain. K. (2006). Children’s reading comprehension: The role of working memory in normal and impaired development. In: *Working Memory and Education* (pp. 61-91). Amsterdam: Academic Press.

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

Chase, W. G., & Ericsson, K.A. (1982). Skill and Working Memory. In: Gordon H. Bower (Ed). Psychology of Learning and Motivation, (Vol. 16, pp. 1-58). New York: Academic Press.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user’s guide. *Psychonomic Bulletin & Review, 12,* 769–786.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*(2), 163.

Cowan, N. (1995). Attention and memory: An integrated framework. Oxford, England: Oxford University Press.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory* (pp. 28–61). Cambridge, U.K.: Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 97–185.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26* (2), 197-223.

Cowan, N. (2015). Second-language use, theories of working memory, and the Vennian mind. In Z. Wen, M.B. Mota, & A. McNeill (eds.), *Working Memory in Second Language Acquisition and Processing.* Bristol, UK: Multilingual Matters. (pp. 29-40)

Cowan, N., Donnell, K., & Saults, J. S. (2013). A list-length constraint on incidental item-to-item associations. *Psychonomic Bulletin & Review*, *20*(6), 1253-1258.

Cowan, N., Fristoe, N. M., Elliot, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition, 34,* 1754-1768.

Cowan, N., Towse, J. N., Hamilton, Z., Saults, J. S., Elliot, E. M., Lacey, J. F., Moreno, M.V., & Hitch, G. J. (2003). Children's working-memory processes: A response timing analysis. *Journal of Experimental Psychology: General, 132,*113–132.

Dahlin, K. (2011). Effects of working memory training on reading in children with special needs. *Reading and Writing, 24*, 479–491.

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

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(4), 561.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, *25*, 1–18.

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

Delaney, P. F., & Ericsson, K. A. (2016). Long-term working memory and transient storage in reading comprehension: What is the evidence? Comment on Foroughi, Werner, Barragán, and Boehm-Davis (2015). *Journal of Experimental Psychology: General, 145*(10), 1406-1409.

Engle, R. W., & Kane, M. J. (2003). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.). Psychology of Learning and Motivation (Vol. 44, pp. 145–199). NY: Elsevier.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of Working Memory* (pp. 28–61). Cambridge, U.K.: Cambridge University Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

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

Everitt, B. S., Landau, S., & Leese, M. (2001). *Cluster analysis*. London: Arnold.

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*(1), 11-15.

Foroughi, C. K., Barragán, D., & Boehm-Davis, D. A. (2016). Interrupted reading and working memory capacity. *Journal of Applied Research in Memory and Cognition*, *5*(4), 395-400.

Foroughi, C. K., Werner, N. E., Barragán, D., & Boehm-Davis, D. A. (2015). Interruptions disrupt reading comprehension. *Journal of Experimental Psychology: General*, *144*(3), 704.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition*, *43*(2), 226-236.

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A. M. (2006). Working memory in children with reading disabilities. *Journal of experimental child psychology*, *93*(3), 265-281.

Gathercole, S. E., & Baddeley, A. D. (1993a). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, *8*(3), 259-272.

Gathercole, S. E., & Baddeley, A.D. (1993b). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. In P.E. Morris & M.A. Conway (Eds.), *The Psychology of Memory, Vol. II*. Aldershot: Edward Elgar Publishing Company, pp. 192-205.

Gathercole, S. E., Brown, L., & Pickering, S. J. (2003). Working memory assessments at school entry as longitudinal predictors of National Curriculum attainment levels. *Educational and Child Psychology, 20*, 109-122.

Gathercole, S. E., Lamont, E., & Alloway, T. P. (2006). Working memory in the classroom. *Working memory and education*, 219-240.

Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Educational Psychology*, *70*(2), 177-194.

Gathercole, S., & Pickering, S. (2001). Research Section: Working memory deficits in children with special educational needs. *British Journal of Special Education*, *28*(2), 89-97.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of experimental child psychology*, *88*(2), 121-151.

Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition*, *36*(1), 93-103.

Healey, M. K., & Miyake, A. (2009). The role of attention during retrieval in working-memory span: A dual-task study. *The Quarterly Journal of Experimental Psychology*, *62*(4), 733-745.

Holmes, J. (2012). Working memory and learning difficulties. *Dyslexia Review, 23*, 7-10.

Jackson, D. L. (2001). Sample size and the number of parameter estimates in maximum likelihood confirmatory factor analysis: A Monte Carlo investigation. *Structural Equation Modeling*, 8, 205–223.

Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N: q hypothesis. *Structural Equation Modeling*, *10*(1), 128-141.

Jöreskog, K. G. (1993). "Testing structural equation models". In Bollen, Kenneth A.; Long, J. Scott. Testing Structural Equation Models. Sage Publications: pp. 294-316.

Jurden, F. H. (1995). Individual differences in working memory and complex cognition. *Journal of Educational Psychology*, *87*(1), 93.

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

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension.* Newton, MA: Allyn & Bacon.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*(1), 103-109.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130(2),* 169–183.

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when an experience-sampling study of working memory and executive control in daily life. *Psychological Science*, *18*(7), 614-621.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637-671.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working-memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*, 189–217.

Kane, M. J., Poole, B. J., Tuholski, S. W., & Engle, R. W. (2006). Working memory capacity and the top-down control of visual search: Exploring the boundaries of ‘‘executive attention”. *Journal of Experiment al Psychology: Learning, Memory, and Cognition*, *32*, 749–777.

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*(6018), 772-775.

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. In B. Ross (Ed.). Psychology of Learning and Motivation, (Vol. 61, pp. 237-284). NY: Elsevier

Karpicke, J. D., & Roediger, H. L. (2007). Repeated retrieval during learning is the key to long term retention. *Journal of Memory and Language*, *57*(2), 151-162.

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. New York: Wiley.

Kintsch, W. (1974). *The representation of meaning in memory.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*, 163-182.

Kintsch, W. (1982). Memory for text. *Advances in Psychology: Discourse Processing*, *8*, 186-204.

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

Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. In N. Onrau & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (Vol. 5, pp. 1270-1328). Newark, DE: International Reading Association.

Kintsch, W., Patel, V. L., & Ericsson, K. A. (1999). The role of long-term working memory in text comprehension. *Psychologia*, *42*(4), 186-198.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363.

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford Publications.

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6), 1118.

Levy, B. A. (1978). Speech analysis during sentence processing: Reading and listening. *Visible Language*, *12*, 81–101.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, *9*(2), 151-173.

Locascio, G., Mahone, E. M., Eason, S., & Cutting, L. (2010). Executive dysfunction among children with reading comprehension deficits. *Journal of Learning Disabilities, 43(5),* 441-454.

Magliano, J. P., Millis, K., Ozurub, Y., & McNamara, D. S. (2007). A multidimensional framework to evaluate reading assessment tools. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theory, interventions, and technologies* (pp. 107–136). Mahwah, NJ: Erlbaum.

McDaniel, M. A., Thomas, R. C., Agarwal, P. K., McDermott, K. B., & Roediger, H. L. (2013). Quizzing in middle-school science: Successful transfer performance on classroom exams. *Applied Cognitive Psychology*, *27*(3), 360-372.

McVay, J. C., & Kane, M. J. (2010). Adrift in the stream of thought: The effects of mind wandering on executive control and working memory capacity. In *Handbook of Individual Differences in Cognition* (pp. 321-334). Springer New York.

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*, *141*(2), 302.

McVay, J. C., Kane, M. J., & Kwapil, T. R. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind wandering across controlled and ecological contexts. *Psychonomic Bulletin & Review*, *16*(5), 857-863.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *45*, 85-86.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). Plans and the structure of behavior. *Operational Research Quarterly*, *19*(3), 338-340.

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press.

Mogle, J. A., Lovett, B. J., Stawski, R. S., & Sliwinski, M. J. (2008). What's so special about working memory? An examination of the relationship among working memory, secondary memory, and fluid intelligence. *Psychological Science, 19*, 1071–1077.

Moscovitch, M. (1992). Memory and working with memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience, 4*, 257–267.

Moulin, C. J., & Gathercole, S. E. (2007). Memory changes across the lifespan. In Cohen, G., & Conway, M. A. (Eds.), *Memory in the real world*. Psychology Press.

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory*, *10*(5-6), 389-395.

Nevo, E., & Breznitz, Z. (2011). Assessment of working memory components at 6years of age as predictors of reading achievements a year later. *Journal of Experimental Child Psychology*, *109*(1), 73-90.

Nevo, E., & Breznitz, Z. (2013). The development of working memory from kindergarten to first grade in children with different decoding skills. *Journal of experimental child psychology*, *114*(2), 217-228.

Oakhill, J., Cain, K., & Yuill, N. (1998). Individual differences in children's comprehension skill: Toward an integrated model. *Reading and Spelling: Development and Disorders*, 343-367.

Pearson, P. D., & Hamm, D. N. (2005). The assessment of reading comprehension: A review of practices--Past, present, and future. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 13-69. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Poole, B. J., & Kane, M. J. (2009). Working memory capacity predicts the executive control of visual search among distractors: The influence of sustained and selective attention. *Quarterly Journal of Experimental Psychology*, *62*, 1430–1454.

Roediger, H. L., & Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*, 181–210.

Roediger, H. L., & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255.

Roediger, H. L., & Nestojko, J. F. (2015). The relative benefits of studying and testing on long-term retention. In J. G. W. Raaijmakers, A. H. Criss, R. L. Goldstone, & M. Styvers (Eds.), *Cognitive modeling in perception and memory: A festschrift for Richard M. Shiffrin* (pp. 99-111). New York: Psychology Press

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General, 126*, 211–227.

Sesma, H. W., Mahone, E. M., Levine, T., Eason, S. H., & Cutting, L. E. (2009). The contribution of executive skills to reading comprehension. *Child Neuropsychology*, *15*(3), 232-246.

Raaijmakers, J. G., & Shiffrin, R. M. (1992). Models for recall and recognition. *Annual review of psychology*, *43*(1), 205-234.

Spillers, G.J. & Unsworth, N. (2011). Variation in working memory capacity and temporal-contextual retrieval from episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1532-1539.

Spillers, G. J., Brewer, G. A., & Unsworth, N. (2012). Working memory and information processing. *Encyclopedia of the Sciences of Learning.*

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643.

Swanson, H. L. (1999). What develops in working memory? A life span perspective. *Developmental Psychology*, *35*(4), 986.

Swanson, H.L. (2003). Age-related differences in learning disabled and skilled readers' working memory. *Journal of Experimental Child Psychology, 85*, 1-31.

Swanson, H.L. & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for math disabilities. *Journal of Education Psychology*, *96*, 471-491.

Swanson, H. L., Howard, C. B., & Saez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities*, *39*(3), 252-269.

Swanson, H.L. & Saez, L. (2003). Memory difficulties in children and adults with learning disabilities. In H.L. Swanson, S. Graham, & K.R. Harris (Eds.), *Handbook of Learning Disabilities*, pp. 182-198. New York: Guilford Press.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, *5*(4), 381-391.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent*? Journal of Memory and Language, 28*, 127–154.

Unsworth, N. (2009). Variation in working memory capacity, fluid intelligence, and episodic recall: A latent variable examination of differences in the dynamics of free recall. *Memory & Cognition, 37*, 837-849.

Unsworth, N. (2010). On the division of working memory and long-term memory and their relation to intelligence: A latent variable analysis. *Acta Psychologica, 134*, 16–28.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2009). There's more to the working memory-fluid intelligence relationship than just secondary memory. *Psychonomic Bulletin & Review, 16*, 931–937.

Unsworth, N., & Engle, R. W. (2007a). On the division of short-term and working memory: An examination of simple and complex spans and their relation to higher-order abilities. *Psychological Bulletin, 133*, 1038–1066.

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

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology, 71*, 1-26.

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

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 832.

Unsworth, N., McMillan, B.D., Brewer, G.A., & Spillers, G.J. (2013). Individual differences in everyday retrospective memory failures. *Journal of Applied Research in Memory and Cognition, 2,* 7-13.

Unsworth, N., Miller, J. D., Lakey, C. E., Young, D. L., Meeks, J. T., Campbell, W. K., & Goodie, A. S. (2009). Exploring the relations among executive functions, fluid intelligence, and personality. *Journal of Individual Differences*, *30* 4), 194.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent variable analysis of the relationship between processing and storage. *Memory, 17,* 635–654.

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention, Memory, or Both? A direct test of the dual-component model. *Journal of Memory and Language, 62*, 392–406.

Unsworth, N., Spillers, G.J., & Brewer, G.A. (2012). Working memory capacity and retrieval limitations from long-term memory: An examination of differences in accessibility. *Quarterly Journal of Experimental Psychology, 65*, 2397-2410.

van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *The Journal of General Psychology*, *132*(4), 408-428.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models an evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, *73*(6), 913-934.

Woodman, G. F., Vogel, E. K., & Luck, S. J. (2001). Attention is not unitary. *Behavioral and Brain Sciences*, *24*(01), 153-154.

Zaromb, F. M., & Roediger, H. L. (2010). The testing effect in free recall is associated with enhanced organizational processes. *Memory & Cognition, 38*, 995 –10.