Spring 5-10-2013

The Influence of Stereotype on Maintenance and Retrieval Errors: Does Working Memory Capacity Matter?

Leslie R. Knuycky
Georgia State University

THE INFLUENCE OF STEREOTYPES ON MAINTENANCE AND RETRIEVAL ERRORS: DOES WORKING MEMORY CAPACITY MATTER?

by

LESLIE RIDDICK KNUYCKY

Under the Direction of Heather Kleider, Ph. D.

ABSTRACT

I explored the influence of stereotypes on performance in cognitive tasks as a function of individual differences in working memory capacity (WMC) in a multi-part study. First, I established that low and high WMC persons maintain equivalent knowledge of common racial stereotypes. Next, I tested whether stereotype-based responses in cognitive tasks that require controlled processing are influenced by individual differences in WMC. Given that stereotypical associations are automatic and cognitively efficient, I predicted that without sufficient resources to suppress these associations, persons with low relative to high WMC will be more susceptible to the influence of stereotype-consistent errors on tasks which have been demonstrated to induce performance differences in low and high WMC persons (Unsworth & Engle, 2007). Engaging WMC is not required in all cognitive tasks; thus, low and high WMC persons were not expected to perform differently on tasks that rely on more automatic processes.

Results provided general support for predictions as persons with more inherently limited cognitive resources committed a higher number of stereotype-consistent errors when performing a maintenance task and accurately recalled fewer stereotype-consistent words when performing a retrieval task. However, persons completing inhibition and familiarity tasks, which are methodologically similar to the maintenance and retrieval tasks but involve less controlled cognitive processes, performed similarly regardless of WMC.

INDEX WORDS: Working memory capacity, Maintenance and retrieval, Stereotypes

THE INFLUENCE OF STEREOTYPES ON MAINTENANCE AND RETRIEVAL ERRORS: DOES WORKING MEMORY CAPACITY MATTER?

By

LESLIE RIDDICK KNUYCKY

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2013

THE INFLUENCE OF STEREOTYPES ON MAINTENANCE AND RETIEVAL ERRORS:

DOES WORKING MEMORY CAPACITY MATTER?

By

LESLIE RIDDICK KNUYCKY

Committee Cahir: Dr. Heather Kleider

Committee: Dr. Sarah Brosnan

Dr. Gwen Frishkoff

Dr. David Washburn

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2013

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1:

INTRODUCTION

1.1 Overview

Knowledge of common stereotypes is ubiquitous and affects everyday actions and reasoning. Persons, regardless of prejudice, have demonstrated equivalent awareness of racial stereotypes (Devine, 1989; Stewart, Weeks, & Lupfer, 2003). Stereotype activation is automatic (Greenwald, McGhee, & Schwartz, 1998; Payne, Lambert, & Jacoby, 2002; Stewart, Weeks, & Lupfer, 2003). Strong stereotypical associations are capable of yielding unintentional evaluation which can influence even deliberative decision making (Oliver, 1999; Oliver & Fonash, 2002; Jones & Kaplan, 2003). Making unbiased decisions that are not influenced by stereotypes requires sufficient working memory capacity (Macrae, Milne, & Bodenhausen, 1994). Thus, it seems that the cognitions of persons with limited cognitive resources (i.e., working memory capacity; WMC) may be more susceptible to the influence of stereotypes. Moreover, greater WMC should be protective against stereotypical bias in decision making. Surprisingly though, extant empirical findings have yet to link individual WMC with stereotypically biased evaluations on cognitive tasks.

A framework put forth by Unsworth & Engle (2007) suggests that specific cognitive tasks (i.e., those requiring active maintenance of information, or retrieval into memory) highlight individual differences in WMC as individuals with a low, compared to high, WMC commit more task errors Thus, if WMC *ever* impacts the influence of stereotypes on decision making, the effect is most likely to emerge in maintenance and retrieval tasks under similar methodological conditions as those already used to establish performance differences between high and low WMC persons.

In the present study, it was my aim to investigate whether individual differences in working memory differentially affect one's reliance on stereotype use during specific cognitive tasks. I posited that persons with more inherently limited cognitive resources would commit a higher number of stereotype-consistent errors when performing maintenance and retrieval tasks. However, not all cognitive tasks induce performance differences between low and high WMC persons (Unsworth & Engle, 2007). Thus, persons completing inhibition and familiarity tasks that are methodologically similar to the maintenance and retrieval tasks but involve different cognitive processes are expected to perform similarly regardless of WMC. In summation, it was my intent to examine the gap in the literature linking WMC to the undue influence of stereotypes. It was my belief that this link has yet to be established due to the nature of the cognitive tasks put forth in experimental protocols (e.g., inhibition tasks). Thus, I proposed a five-part investigation wherein I attempted to 1) establish whether low and high WMC persons maintain similar knowledge of common racial stereotypes, 2-3) test the prediction that WMC affects the influence of stereotype use on maintenance tasks (2) but not inhibition tasks (3), and 4-5) test my prediction that WMC affects the influence of stereotype use on retrieval (4) but not familiarity tasks (5).

To build and support my thesis, I will briefly review the evolution of working memory and working memory capacity; discuss research demonstrating performance differences as a function of individual capacity; qualify when and how stereotypes are expected to affect decision making; and finally, specify the hypotheses and methods used in the current investigation.

1.2 Working Memory (WM)

The theoretical construct, “working memory,” was proposed by Baddeley and Hitch, 1974 to account for problems that Atkinson and Shiffrin’s (1968) influential Modal Model of memory failed to explain. The Modal Model of memory proposed three stages of memory, or memory subsystems: Sensory memory, short-term memory, and long-term memory. These systems were thought to differ in capacity and duration, as well as other features, such as the main modality for encoding (e.g., visual, auditory, or semantic "codes"). According to this model, sensory memory holds a vast amount of raw sensory information for up to 10 or 15 seconds (Sperling, 1960). Short-term memory (STM) lasts for 20-30 seconds, is extremely limited (Miller, 1956; Cohen REFS), and can support semantic, as well as visual or phonological representations (REFS). Long-term memory (LTM) is virtually unlimited (e.g., new words or facts can be added without limit), includes associations within and across modalities, and has a duration of anywhere from a few minutes to 50 or more years. Importantly, the Modal Model proposes that these three memory processes are strictly sequential: information must be processed in short-term memory in order to be encoded in long-term memory.

Two key findings presented difficulties for the Modal Model. First, some individuals with severe cognitive impairments, such as patient K.F., show intact LTM (i.e., normal encoding and retrieval of information in LTM) in the absence of fully functioning STM (Shallice & Warrington, 1970). Thus, it appears that a fully intact STM is not necessary for encoding of information in LTM. Although this observation is not inconsistent with the Modal Model, neither is it predicted by the model. Therefore, either the Modal Model has inadequate definitions of STM or LTM (or both), or it is missing a key construct. Second, studies of dual task performance have shown that, under some circumstances, people can perform two tasks at once, without major performance decrements on either task. For example, Baddeley and Hitch (1974) presented participants a sequence of four single-digit numbers and asked them to keep them in memory while reading a passage. The fact that subjects could comprehend the reading and successfully recall the numbers suggested that memory was capable of both the storage and complex manipulation of information. These results led Baddelely to propose the name *working memory* for short-term processes replacing the previous name of short-term memory. Baddeley defined working memory as a limited capacity system, "that provides temporary storage and manipulation of the information necessary for such complex tasks as language comprehension, learning, and reasoning" (Baddeley, 1992, p. 556).

In the Baddeley model, working memory comprises four components: central executive, phonological loop, visuospatial sketchpad, and (in later versions of Baddeley's model; Baddeley, 2000b) the episodic buffer. The central executive pulls and integrates information across the subsystems; the phonological loop, visuospatial sketchpad, and episodic buffer. The phonological loop holds and processes verbal and auditory information while the visuospatial sketchpad holds and processes visual and spatial information. Rehearsal in the phonological loop prevents loss of information by keeping information active. Finally, the episodic buffer moves information to and from episodic long-term memory.

There is considerable support for Baddeley's working memory model. For example, Baddeley, Lewis, and Vallar, 1984 showed that a list of short words is recalled more accurately than an equally long list of longer words (word length effect). The word-length effect occurs because the phonological loop has limited capacity (like STM in the Modal Model) and because information is actively maintained in the phonological loop through rehearsal and shorter words can be rehearsed at a faster rate, leading to more robust encoding. Articulatory suppression also provides support for this model. Rehearsal of information in the phonological loop engages sub-vocal or articulatory processing (i.e., inner speech). Therefore, if a person continuously a word or word-like sequence (e.g., "*the, the, the…*") during encoding, memory recall is greatly impaired (Baddeley, et al., 1984). Moreover, the word list effect disappears under articulatory suppression. Further, whereas articulatory suppression hinders recall of spoken-word lists, it benefits performance of nonverbal tasks. Specifically, Brandimonite, Hitch, and Bishop (1992) showed participants a picture of an object (e.g., piece of candy in a wrapper) and then presented them with a picture of part of the object (e.g., end of wrapper). Without the pictures present participants were asked to subtract the second picture from the first and then describe the resulting picture (i.e., a fish in the current example). Half of the participants repeated *la, la, la* (articulatory suppression) during the task. Surprisingly, the participants in the articulatory suppression group outperformed the other group on the subtraction task. This result occurred because the articulatory suppression group was forced to encode the object as a picture into the visuospatial sketchpad instead of as a word (i.e., the object name) into the occupied phonological loop. Thus, manipulating the picture for the subtraction task was made easier as this group started with a visual image compared to a verbal code.

1.3 Working Memory Capacity (WMC) and Individual Differences: In the Beginning

All models of working memory have assumed that the capacity was finite.[1] Furthermore, researchers recognized that the amount of available capacity affects subsequent processing. Originally, however, it was suggested that skill or efficiency was what created or limited

---

[1] See Morrison & Chein, 2011, for a recently developing discussion regarding the effectiveness of capacity enhancing techniques

remaining capacity (Daneman & Carpenter, 1980). That is, the thesis suggested that the capacity limits of working memory were fixed across all individuals. Differences in performance were the result of efficiency within the tested domain. Under this framework, the task of reading would require less capacity for an efficient, compared to an inefficient, reader leaving additional capacity for an alternative task. However, the efficient reader may be inefficient at solving arithmetic problems and thus use more capacity for the latter task leaving less residual capacity. It was this theory that drove Daneman and Carpenter to create the reading span task (described below) wherein they found high correlations between reading span and various reading comprehension measures. Importantly, they did not find a significant correlation between a simple word span test[2] and the comprehension measures. The authors argued that the reading span measure reflects the amount of working memory capacity remaining after allocating the required resources (which varied depending on participant efficiency) for the processing task. Furthermore, they suggested that the resulting span score would vary as a function of the background processing task. Thus, in order to successfully predict reading comprehension the span task must include a reading processing task.

Turner and Engle (1989) offered a different explanation for the Daneman and Carpenter (1980) findings. They suggested that working memory capacity may be affecting reading ability instead of reading ability affecting the capacity remaining above and beyond the processing task. Stated differently, all persons may not be working with the same fixed capacity. Instead, a larger working memory capacity may have allowed readers to process and comprehend the material while concurrently holding several words in memory, whereas a smaller working memory

[2] Simple word span tasks orally present participants with a series of 2-7 single-syllable common nouns that they must recall in presentation order. List length increases until a participant fails to accurately recall all words for 3 out of 3 lists. The participant's word span was considered to be the level wherein they accurately recalled all words for 2 out of 3 lists.

capacity may have limited both the reader's comprehension of the material and the number of words that they could keep in memory. Furthermore, Turner and Engle suggested that the capacity limit was not domain specific. That is, a large working memory capacity should benefit the processer regardless of the task being performed. To test this hypothesis the authors collected four complex span and two simple span measures from participants and then conducted a reading comprehension test. The complex span measures included: 1) the traditional reading span task used in the Daneman and Carpenter (1980) study, 2) an alternative reading span task wherein the to-be-remembered stimuli were digits instead of words, 3) an operation span task that required participants to solve arithmetic problems while holding words or 4) digits in memory. The simple span measures required participants to hold increasingly larger lists of words or digits in memory and recall as many as possible in the order that were presented, these measures are referred to simply as word or digit span respectively. The idea was that if reading ability affected the remaining capacity and if capacity is domain specific, as was suggested by Daneman and Carpenter, then individual reading comprehension should correlate with the two reading span scores but not with either of the operation span scores. Alternatively, if capacity affects reading ability and if capacity is not domain specific, then reading comprehension should correlate with all four complex span scores. The simple span scores were not expected to predict reading comprehension. Results revealed a significant correlation between reading comprehension and all four complex span scores. Simple span scores did not correlate with reading comprehension. Furthermore, the operation word span task reliably predicted reading ability even after the variance accounted for by reading span was removed. Also, it is worth noting that the four span tasks correlated significantly with each other (r=.38-.58). Taken together, findings suggest that, as proposed by Turner and Engle, WMC is not domain-specific and that capacity affects reading comprehension instead of reading ability affecting the amount of capacity remaining for alternative tasks.

In 1992 Just and Carpenter proposed the capacity theory of comprehension. The theory explained capacity as the maximum amount of activation available in working memory to support storage and processing as both functions draw from the same pool of resources. Furthermore, the researchers suggested that capacity constrains performance on cognitive tasks, and it does so more for some people than for others. The authors offered individual differences in working memory capacity as the cognitive mechanism underlying results from earlier studies that had previously been explained by differences in cognitive architecture. Specifically, Ferreira and Clifton (1986) presented participants with sentences wherein the head noun was either animate, and therefore could act or be acted upon, or inanimate, and thus, could not be responsible for acting. Animacy was considered a pragmatic cue. Example sentences include: "The *defendant* that was examined by the lawyer shocked the jury," versus "The *evidence* that was examined by the lawyer shocked the jury." Furthermore, half of the sentences were reduced such that disambiguating information following the initial noun (i.e., 'that was') was removed. The expanded relative clause was considered a syntactic cue. Utilizing the pragmatic and syntactic cues concurrently should result in equal comprehension of all sentence-types with the exception of the reduced/animate sentences. Specifically, the meaning of reduced sentences with an animate head noun would remain unclear until the entire sentence was read. For example, in the sentence, "The defendant examined by the lawyer shocked the jury," it is unclear whether the defendant *was* examining or *was being* examined until the phrase "by the lawyer" is read. The researchers measured participant eye gaze while the sentences were read. The results revealed that participant gaze duration was longer for the reduced compared to the unreduced sentences but animacy had no effect on eye gaze. The authors suggested that the architecture of the system is such that syntactic processing is encapsulated (i.e., its activities and outputs are unaffected by information that exists elsewhere in the system).

Just and Carpenter (1992) rejected the theory regarding encapsulated cognitive architecture and instead postulated that interaction between processing systems (i.e., pragmatic and syntactic) requires capacity. Capacity, they theorized, was dependent upon the individual. To test this idea, the authors repeated Ferreira and Clifton's (1986) experiment but this time looked at high compared to low span individuals separately. Replicating previous results, reading times were faster for the unreduced compared to the reduced sentences. However, the primary finding of interest was the significant difference in eye gaze duration for high- compared to low-span participants when reading the inanimate sentences. The high-span individuals spent significantly less time reading the inanimate compared to the animate sentences whereas animacy had no effect on gaze duration for low-span individuals. The authors explained this finding as the result of high-span individuals utilizing the pragmatic cue of animacy. Thus, the authors concluded that interaction between processes was possible (i.e., processes were not modular and embedded), but that such interaction required capacity. It is important to note that the described results pertain to the first pass reading only and that during later readings both high- and low-span individual's gaze duration decreased for inanimate compared to animate sentences. This indicates that all individuals can ultimately make use of the pragmatic cue but only high-span individuals do so on a first pass reading.

MacDonald, Just and Carpenter (1992) offered another instance wherein individual differences in capacity explained previously conflicting theories. Specifically, researchers had been debating how individuals may handle ambiguities when reading text. On one side of the argument it was thought that immediately upon encountering an ambiguity the reader chose a single interpretation (Just and Carpenter, 1987); conversely, others argued that the reader retained both options in mind until the ambiguity was later resolved in the text (e.g., Kurtzman, 1985). MacDonald and colleagues proposed that both camps were valid but under varying circumstances. They suggested that all readers activate multiple representations upon encountering an in-text ambiguity. The representations are weighted based on its frequency, complexity, and plausibility. Differences emerge between high- and low-span readers as they continue to read. That is, low-span individuals lack sufficient capacity to process incoming material while continuing to store multiple interpretations in mind; thus, they drop the less common interpretation and move forward with a single interpretation. High-span individuals continue to maintain both interpretations for a significantly longer duration. The authors found support for their position with what may initially be a counterintuitive result. High-span readers took longer to read ambiguous compared to unambiguous sentences, particularly at the point wherein the ambiguity is resolved. The effect of sentence ambiguity was generally unreliable for low-span readers as they read both sentence types equally fast. The authors explained the somewhat surprising slowdown in processing for the high-span readers as the cost of storing multiple interpretations of the ambiguous item. However, the payoff for high-span individuals is evidenced with tests of comprehension, particularly when the accurate interpretation for the ambiguous item is the less common interpretation, in these instances high-span individuals perform well whereas low-span individuals often score at or near chance.

This brief depiction of the evolution of working memory capacity chronicles theories and research starting from the point at which all individuals were assumed to be operating with the same three levels of memory (i.e., modal model of memory). Presently the thesis of a finite capacity that varies across individuals and predicts task performance and abilities is widely accepted. Furthermore, there is currently substantial support for controlled attention views of WMC (Kane, Conway, Hambrick, & Engle, 2007). I will now turn to the ways in which individual working memory capacity is commonly measured and current WMC theories and research.

1.4 Measuring Span

A variety of tasks have been used to measure individuals working memory capacity (WMC). In each task the individual is forced to draw on their resources from working memory in order to process stimuli and simultaneously store information for retrieval at a later point. The resulting span score reflects how successful the individual is able to negotiate both parts of the task. Broadly speaking, span tasks measure individuals' ability to multitask. Some of the more common span tasks are explained in more detail below.

1.4.1 Reading Span

The reading span task (Daneman & Carpenter, 1980) requires participants to read a set of unrelated sentences and hold the last word of each sentence in memory for later recall. Reading span is defined as the maximum set size for which a participant can accurately recall all of the final words for at least three of five sets. On average, college students' reading spans range between 2 and 5.5 with high span individuals accurately recalling all words in sets of four or more, medium spans accurately recalling all words of in sets of three to three and a half, and low span individuals accurately recalling all words in sets less than three.

1.4.2 Operation Span

The operation span task (Unsworth, Heitz, Schrock, & Engle, 2005) requires participants to determine the validity of two-operation math sentences (e.g., (4+2)/3=6; true or false). After each math sentence a to-be-remembered letter is presented. Following a sequence of 3-7 math-letter strings, participants are asked to recall the list of letters in the order that they were shown. A total of 75 math-letter strings are presented throughout the duration of the task. Operation span is defined as the number of letters accurately recalled in presentation order.

1.4.3 Symmetry Span

The symmetry span task (Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle, 2004) requires participants to perform a symmetry judgment task. After each judgment a red square is presented within a 4 x 4 matrix. After 2-5 red squares are shown, participants are asked to recall the sequence of red square presentations in the order that they appeared by clicking on the various locations on an empty matrix. A total of 42 symmetry-red block strings are presented throughout the task. Symmetry span is defined as the number of red blocks accurately recalled in the order that they were presented.

1.4.4 Complex Span Tasks

In practice, complex span tasks, like those described above, are the accepted method of assessing WMC. Overall, complex span tasks have shown good reliability and validity (see Unsworth, Redick, Heitz, Broadway, & Engle, 2009). Scores on complex span tasks (or WMC) have been linked to a variety of phenomenon including early onset Alzheimer’s (Rosen, Bergeson, Putnam, Harwell, & Sunderland, 2002), susceptibility to life-event stress (Klein & Boals, 2001), and susceptibility to stereotype threat (Schmader & Johns, 2003). Recently, researchers have found that complex span tasks are predictive of multiple aspects of cognition including recall, processing accuracy, processing time, and fluid intelligence (Unsworth et. al., 2009). Furthermore, the three components that comprise the complex span task account for 69% of the variance in fluid intelligence.

1.5 Working Memory Capacity (WMC) and Individual Differences: Theories

Whereas WMC is related to multiple higher order cognitive processes (e.g., reading and reasoning), working memory is not required for all tasks. Unsworth and Engle (2007) put forth a framework suggesting that performance differences would emerge between low and high working memory persons only under certain conditions. Specifically, they stated that a) working memory is needed in order to override automatic response biases and, b) it fulfills two basic functions, maintenance and retrieval. First, maintenance, or focused attention, occurs in primary memory and is required to keep novel information active, particularly when internal or external distractions are present. Importantly, primary memory is a limited but also flexible capacity. Thus, depending on the task, primary memory may hold more or less information active at a given time. For example, research by Cowan (2001) has demonstrated that during a word memory task where the goal is to retain as many representations as possible, primary memory holds 4 +/- 1 items. When the task goal changes and requires strong control to overcome automatic responses with success dependent upon goal maintenance primary memory capacity can shrink to one representation. Second, because primary memory is limited in capacity, displaced items must be retrieved from secondary memory which is responsible for the storage of both task relevant and task irrelevant items. Retrieval of task relevant items is guided by a discrimination process through the use of cues. Thus, the authors suggest that high working memory persons are better able to 1) maintain representations in primary memory and, 2) conduct controlled searches in secondary memory.

Evidence supporting the aforementioned framework comes from a variety of research. Beginning with individual differences due to maintenance, take for example studies using the antisaccade task. The task requires that participants make a response either toward (prosaccade) or away from (antisaccade) a flashing cue. The automatic and task-required responses align in the prosaccade trials, thus, there is no need to override the automatic response or to maintain the task goal in memory in order to perform well. High and low working memory participants performed equally well on the prosaccade trials (Kane, Bleckley, Conway & Engle, 2001). However, in the antisaccade trials, the automatic and task-required responses are misaligned; therefore, control to override the automatic response tendency and goal maintenance is imperative. High working memory participants completed the antisaccade trials faster and with fewer errors than the low working memory participants (Kane et al., 2001).

A second example of a task requiring goal maintenance for quality performance comes from the Stroop task. The Stroop task consists of color names (e.g., GREEN, RED, BLUE) presented in colored ink and requires participants to name the ink color. Sometimes the ink color is congruent with the color name (e.g., GREEN printed in green ink) other times the ink color is incongruent with the color name (e.g., GREEN printed in red ink). When trials comprise either all congruent or all incongruent words, goal maintenance is unnecessary because the previous trial reinforces the goal of the current trial. However, when some trials are congruent and some incongruent, goal maintenance becomes essential. It is in these intermixed trials that the Unsworth and Engle (2007) framework predicts that performance for high, compared to low, WMC persons should diverge. That is precisely what researchers have found. Kane and Engle

(2003) presented participants with either incongruent-only or intermixed Stroop trials. In the incongruent-only condition there was a negligible difference in error rate as a function of WMC. Conversely, in the intermixed condition, low working memory persons made significantly more reading errors than high working memory persons. Kane and Engle suggested that the result was due to low WMC persons' inability to maintain the task goal and instead reverting to the automatic reading response.

The second part of the Unsworth and Engle (2007) framework suggests that individual differences should emerge during retrieval. The authors argue that a controlled search into secondary memory requires the use of contextual cues in order to discriminate between task relevant and task irrelevant information. These cues are organized into a hierarchy and change at varying speeds (e.g., Unsworth & Engle, 2006). Using the example of learning word lists, at the top of the hierarchy and unlikely to change during the learning period are global cues (e.g., the learning location). Next, there are cues related to each individual list. Finally, at the bottom of the hierarchy are rapidly changing cues associated with each individual word. High WMC persons are said to encode items into secondary memory using cues lower on the hierarchy than low WMC persons. Support for this thesis comes from research using recall tasks. Specifically, low WMC, compared to high WMC, persons have a tendency to recall fewer target items and a greater number of previous list intrusions as the list recall task progresses beyond the first list (Kane & Engle, 2000). The authors argued that because low WMC persons use contextual cues higher on the hierarchy, they actually create a noisier search set that includes items from previously learned lists as well as target items from the present list, thus, hindering their performance.

It is interesting that in a variant of the recall task, researchers sometimes instructed participants to forget previously learned lists. Later, participants were unexpectedly asked to recall words from the forgotten list. Under these conditions, high WMC persons recalled fewer words than low WMC persons (Delaney & Sahakyan, 2007). The authors suggested that the finding was the result of the discrepancy between the context at encoding and retrieval. This mismatch resulted in high WMC using contextual cues to delimit the search to the point of excluding the target items. However, as the low WMC persons use less specific contextual cues to confine their search, they maintained the 'forgotten' words in their search sets.

Recognition studies provide additional evidence for individual differences in retrieval processes. Performance in recognition tasks is the result of two separate mechanisms: familiarity which is automatic and recollection which requires control (see Yonelinas, 2002). Only the latter is subject to performance differences as a function of individual WMC. For example, Bunting, Conway, and Heitz (2004) presented participants with statements linking two concepts (e.g., the lawyer is in the park). The authors manipulated the number of links that were associated with each concept. During a recognition test, low and high WMC persons responded equally to concepts with only a single link, as this response was based on a feeling of familiarity. However, high WMC persons responded more quickly than low WMC persons to concepts with multiple links; because a response required a controlled search into secondary memory, high WMC persons' relatively limited search criterion led to a speedier response.

In summation, the framework put forth by Unsworth and Engle (2007) asserts that individual differences in WMC emerge when control is required in order to 1) override automatic response biases on tasks that require goal maintenance in primary memory, or 2) conduct a delimited controlled search into secondary memory.

1.6 Diverging Theories

Other theories have been offered as explanations for many of the WMC results. Specifically, some researchers suggest that inhibition alone, or the ability to stop an automatic response is the key mechanism underlying a variety of cognitive abilities including WMC (Lustig, May, & Hasher, 2001). According to this framework, it is the high, relative to low, WMC persons' superior ability to inhibit automatic responses that creates individual differences. Alternatively still, others suppose a Binding theory to explain individual differences in WMC (Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007). The Binding theory suggests that it is the strength of the stimulus-response (S-R) relationship that creates the response differences between high and low WMC individuals, with stronger S-R bindings in high, than low, WMC persons.

Redick, Calvo, Gay, and Engle (2011) conducted a series of go/no-go experiments in order to test these diverse theories. Experimental stimuli included the presentation of various consonant letters. The first experiment included two trial blocks. In the first block, participants were required to respond go to the presentation of the letter X and provide no response (i.e., no-go) to all other letter presentations. Block 1 was composed of 80% go-trials. In the second block the instructions were reversed such that participants were to respond go to the presentations of any non-X letter and not respond to the letter X. According to the inhibition theory of WMC, low WMC persons should have a more difficult time withholding (i.e., responding with no response) on the no-go trials than high WMC persons, and this difference as a function of WMC should occur in both blocks of the experiment. The Binding theory of WMC predicts that the greatest differences between high and low WMC individuals should occur in the second block as high WMC persons should be better able to update their S-R relationship. Finally, the maintenance and retrieval framework does not suggest a difference in the performance of high and low working memory persons as the task requires only that participants maintain a single and consistent (within block) rule, therefore primary memory is not significantly taxed. The results revealed a main effect of trial-type such that participants made fewer errors on the go- than no-go trials. The main effect of block was also reliable as performance decreased from the first to the second block. However, there was no effect of WMC or any significant interactions.

Next, the experimenters modified the task slightly. They changed the task instructions to require a go-response to the letters W or M but only when the identity of the target letter had changed from that of the previous target. Thus, a go-response to the letter W was only accurate if the pervious target was the letter M. A target letter with the same identity as the previous target letter was considered a lure in the present study. This minor manipulation should change participants' strategy for a successful performance. Specifically, whereas the first experiment required minimal maintenance, participants must have continually updated the task goal and actively maintained it over time in order to respond accurately in the second experiment. Therefore, the maintenance and retrieval theory predicted significant performance differences between high and low WMC individuals in the modified go/no-go task, particularly in response to trial lures. The results supported these predictions. High and low WMC persons responded with equal accuracy to distractor trials (i.e., non M or W letters, always requiring a no-go response). However, low WMC persons made significantly more errors on the target (i.e., M or W targets requiring a go response) and lure trials (i.e., M or W targets requiring a no-go response). In additional analyses, the authors looked at low compared to high WMC persons' accuracy to lure trials as a function of time between the current lure and the previous target. Stated differently, some lures occurred immediately after a target whereas other lures occurred with 1, 2 or 3 distractor items between the lure and the previous target. If it is the case that low WMC individuals have a more difficult time maintaining the task goal over time than high WMC individuals then accuracy to lure trials should decrease as the number of intervening distractors increase for low WMC persons. This is exactly what the authors found. Low and high WMC persons performed equally well on lure trials occurring immediately after a target trial, thus, the maintenance requirement was relatively minimal. Conversely, low WMC persons' performance declined steadily as the number of intervening distractors (i.e., time) increased, whereas high WMC persons maintained a consistent level of performance regardless of the time delay since the previous target. This final analyses depicting a performance difference for high, compared to low, WMC individuals as a function of the task maintenance requirement provides strong support for the maintenance and retrieval theory of individual differences in WMC and is difficult for the other theories (i.e., Inhibition and Binding) to explain.

1.7 Working Memory Capacity (WMC) and Individual Differences: Decision Making

Kahneman (2003) explained the distinctions between intuition and reasoning. He suggested that intuition is fast, automatic, effortless, associative, and emotional. Conversely, reasoning is slow, controlled, effortful, rule-governed, flexible, and neutral. Thus, cognitive resources are required in order to allow for reasoning, which is more logical and less likely to be influenced by irrelevant circumstances. Researchers have linked individual differences in working memory capacity to differences in decision making. Specifically, high WMC persons tend to make more consistent judgments across scenarios than low WMC persons.

For example, Moore, Clark, and Kane (2008) had participants read moral dilemmas and judge the moral appropriateness of killing one person in order to save others. The dilemmas included scenarios wherein the death of another was: 1) necessary to save a) oneself and others or b) only others; 2) a) inevitable regardless of the subject's actions or b) could be avoided; and 3) a) a direct (i.e., pushing the target) or b) indirect (i.e., push a button that caused the target to fall) result of the subject's actions. The runaway trolley is an example of a typical moral dilemma. In this dilemma there is a trolley barreling down the tracks towards five unaware workers. If the trolley continues on its current path it is certain to kill all of the workers. In the direct-action version of the dilemma the participant must physically throw a heavy man from the train platform onto the tracks in order to stop the trolley and save the workers. In the indirect-action version of the dilemma the participant must push a button that will divert the trolley down a different track where a single workman will be hit and killed but the lives of five workmen will be spared. In each version of the dilemma, the subject's actions lead to the death of one workman which prevents the deaths of five workmen; thus, the outcomes are the same. The directness of the physical intervening action should not logically affect the moral appropriateness of the killing. Importantly, moral appropriateness judgments were not time limited.

Nevertheless, overall, killing to save oneself and others, killing when death was inevitable, and indirect killing were each determined to be more morally appropriate than killing to save only others, killing when death was avoidable and direct killing. Of primary interest for the present discussion was the role of individual WMC on judgments of moral appropriateness. When the death of another was avoidable all persons judged it to be equally morally appropriate, however, when the death was inevitable persons with high, compared to low, WMC found it to more morally appropriate. Additionally, persons with high WMC were found to make more consistent judgments of moral appropriateness than persons with low WMC when the dilemmas required direct action. The authors suggest that these findings were the result of high WMC persons' increased deliberative judgments particularly in the cases requiring direct action in order to kill when the death was inevitable. Supporting this argument, Greene and colleagues (Greene, Nystorm, Engell, Darley, & Cohen, 2004) proposed that moral dilemmas requiring direct action evoke an automatic emotional response against causing harm. However, the biased (and irrational) response can be overridden by cognitive control.

In a similar line of research, authors investigated participants' tendency to engage in counterfactual thinking as a function of WMC (Goldinger, Kleider, Azuma, & Beike, 2003). Counterfactual thinking often occurs after a negative event and involves the generation of alternative event endings. This tendency is spontaneous (Kahneman, 1995) and thus, avoiding it should require effortful suppression. In order to test this prediction the authors first assessed WMC via the operation span task. Next, they presented participants with 8 short stories wherein someone is the victim of an unforeseen accident. Half of the stories manipulate the victims' actions such that a counterfactual alternative is made salient. For example, in one control story, a worker leaves the job after a typical workday (i.e., at 5:30) and is broadsided on the way home. Alternatively, in the counterfactual version of the story, the worker becomes restless on the job and decides to leave early to see a movie (i.e., 4:30) and is broadsided on the way to the theatre. The time that the victim leaves work should have no bearing on the degree to which s/he is determined to be at-fault for the incident or the amount of compensation that s/he is awarded. However, the counterfactual version of the story makes an alternative ending to the event particularly easy to create (i.e., if only s/he had been more responsible and stayed at work until 5:30, then this would not have happened) and thus, the victim appears responsible for the outcome.

During the experimental procedure, participants were presented with an external load. Specifically, participants were shown 6 bisylliabic nonwords that they were to study for later recall. Participants either read case facts or made juror judgments regarding victim compensation, and victim guilt while keeping the nonwords in memory. Results revealed that low WMC persons awarded significantly less compensation to victims and attributed a greater amount of blame to the victim in the counterfactual story condition than in the control conditions. Conversely, high WMC persons were unaffected by the story condition when deciding victim blame and compensation amount.

Each of the aforementioned studies demonstrates the importance of cognitive resources in making consistent and logical decisions that are not biased by irrelevant details.

1.8 Individual Working Memory Capacity (WMC) and New Directions

As reviewed above, WMC varies across individuals and affects judgments and decision making on a variety of tasks. Low, compared to high, WMC persons' cognitive resources are more completely taxed under certain conditions. In such circumstances irrelevant but automatically activated task information is more likely to inform low WMC persons' decision making (e.g., the emotionality of direct versus indirect killing; the counterfactual actions of the victim). What has not yet been tested (with the exception of Kleider, Knuycky, & Cavrak, 2012; see study overview below), is whether high WMC is protective against undue stereotype influence. That is the question under investigation in the present study. Specifically, using Unsworth and Engle's (2007) maintenance and retrieval theory of individual differences in WMC, in the current study I will test whether low, relative, to high WMC persons' differ in their tendency to make stereotype-consistent errors on maintenance and retrieval tasks. As stereotypes are automatically activated (see examples and elaboration below) and require resources to suppress, I posit that low WMC persons will be more susceptible to the influence of stereotypes than their high WMC counterparts.

1.9 Stereotypes are Ubiquitous and Natural

Person perception models suggest that the first step in perceiving someone or something includes categorization (Brewer, 1988; Fiske & Newberg, 1990). Whereas categorization is beneficial in that it is cognitively economic; the process is not without negative consequences. Categories and category memberships allow for prejudgment as "*the category saturates all that it contains with the same ideational and emotional flavor*" (Allport, 1954; pg 21). Categorization is thus both a necessary and sufficient prerequisite to stereotyping. Fiske (2004) defines stereotyping as *applying to an individual one's cognitive expectancies and associations about the group*. Common societal stereotypes are transmitted efficiently through the social and cultural environment with children as young as three demonstrating stereotype awareness (Cramer & Steinwert, 1998). Stereotypes are learned, or over learned, to the point of being automatically activated (i.e., unintentionally, involuntarily, effortlessly, and outside awareness; see examples below), thus, it should be of no surprise that stereotypes often have undue effect on judgment, memory, and decision making.

1.10 Stereotypes are Automatically Activated

Although everyone may not endorse a common societal stereotype, these ingrained associations are activated automatically. For example, Blair and Banaji (1996, Experiment 2) presented participants with stereotypic gender traits followed immediately by a male or female name. Participants were to identify the gender of the name. Findings revealed that participants more quickly identified male than female names after the presentation of stereotypical male traits and more quickly identified female than male names after the presentation of stereotypical female traits. As the time between the trait and name presentation (250 ms) was insufficient to allow for controlled processing, the authors argued that this result was due to automatic stereotype activation. Specifically, the gendered trait served as a prime. Thus, when the trait/name pair was stereotype-consistent response time was fast. Conversely, when the gendered trait was presented prior to a stereotype-inconsistent gendered name, automatically activated expectations were violated and response time suffered.

Researchers have found that racial stereotypes were automatically activated for low and high prejudice persons alike (Stewart, Weeks, & Lupfer, 2003). In this study, participants were shown head and shoulder photos of Black and White men presented concurrently with three behavior statements. The behavior statements depicted either stereotypical or non-stereotypical African American traits (e.g., trait: athletic; behavior statement: often played in pickup basketball games after school). Immediately following the presentation of the photo/behavior statements, participants were given a probed recognition memory test. The recognition test required that participants indicate (yes/no) whether any of four presented words had been included verbatim in the previous behavior statements. The probe word for the stereotypical behavior statement was the implied trait. Participants, regardless of individual prejudice, were more likely to mistakenly say that they remembered seeing the implied trait when the stereotypical behavior statement was paired with a Black than a White face. These finding support the conclusion that, independent of personal convictions, all persons are aware of, and sometimes influenced by, automatically activated stereotype associations.

Researchers capitalized on automatic stereotype activation in an effort to measure implicit prejudices. The result was the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). The IAT presents participants with the category labels ‘pleasant’ and ‘unpleasant’ with one each on the left or right side of the computer screen. Participants are asked to categorize unambiguous words (e.g., ugly or peace) according to the appropriate label as quickly as possible. Next, the labels ‘Black’ and ‘White’ are presented. Participants are asked to categorize ethnic names (e.g., Jamal or Adam) according to the appropriate demographic. After several trials with the single labels, the two labels from the two categories were presented concurrently in either consistent (pleasant and White; unpleasant and Black) or inconsistent (unpleasant and White; pleasant and Black) trials. The reaction times for correct word and name categorizations were averaged for the consistent and inconsistent trials. A faster average reaction time for the consistent trials reflects a strong stereotypical association and is often labeled as reflecting an implicit bias against Black persons.

1.11 The Efficiency of Stereotypes

“As a rule monopolistic categories are easier to form and to hold than are differentiated categories.” (Allport, 1954)

A clever set of studies conducted by Macrae, Milne, and Bodenhausen (1994) demonstrated the cognitive efficiency of stereotypes. In the first experiment, participants completed two tasks concurrently; they formed an impression of a target using trait descriptors while simultaneously monitoring audio prose. The authors manipulated the presentation of the trait descriptors such that half of the participants saw the targets with a stereotype label (e.g., doctor, skinhead) and half did not. The authors reasoned that if the use of stereotypes is energy efficient, then, forming impressions of targets described with a stereotype label (versus targets described without a label) should preserve capacity allowing for enhanced performance on the prose monitoring task. Their hypothesis was supported as participants forming impressions with the aid of a stereotype label (compared to no label) scored higher on a multiple choice questionnaire testing memory for the presented prose. In a second experiment, the authors repeated the study but made one important change; they presented the label subliminally. The results from study one were replicated. In a final study the authors attempted to quantify the resources preserved via stereotype use. In this study the second task was changed (previously the prose monitoring task); participants were now required to monitor for and respond to a beeping noise. Reaction time was measured. Results revealed that stereotype activation (supra- or subliminal) lead participants to respond nearly 230ms faster to the monitoring task (supraliminal stereotype label: 582ms; subliminal stereotype label: 571ms; no stereotype label: 812ms). In combination, these results provide convincing evidence that stereotype use is cognitively efficient.

As stereotyping is energy efficient it is often the default when cognitive resources are limited. One example of stereotypes affecting decision making when controlled processing was precluded comes from Payne, Lambert, and Jacoby (2002). Researchers presented participants with a prime consisting of either a Black or a White face, followed by a picture of either a tool or a weapon and then asked participants to identify the object. Participants primed with the Black face more accurately identified the weapon and misidentified the tool, a tendency that persisted regardless of instructions to avoid using race as a cue. A follow-up study revealed that participants were immediately aware of their object classification errors and were able to provide the correct classification when given a second opportunity, but were unable to suppress the erroneous classification initially (Payne, Shimizu, & Jacoby, 2005). The authors suggested that the robustness of activated stereotypes make it difficult to suppress rapid response decisions. In a separate but related study Payne (2005) found a relationship between cognitive control and stereotype use. Payne used participant scores on weapon identification and word evaluation tasks to calculate a process dissociation procedure (PDP) in order to discern estimates of automatic versus controlled components of performance. The author used this measure to explain individual differences in stereotype-consistent errors and judgments. The results revealed that participants utilizing relatively more control during tasks made fewer stereotype-consistent errors and judgments. This finding demonstrates that when greater executive control is engaged automatic biases decrease.

Stereotype use can be avoided, however, there is a cognitive cost; monitoring stereotypes requires mental resources. Specifically, Richeson and Shelton (2003) investigated whether an interracial interaction was cognitively taxing for high- relative to low-prejudice persons. To that end, researchers presented White participants with an IAT, engaged them in a 2-minute recorded interview with either a same- or other-race experimenter, and finally assigned a Stroop task for completion. Participants with an above-average IAT bias score revealed greater Stroop interference after interacting with the Black than the White interviewer. Participants with a below-average IAT bias score did not exhibit a difference in Stroop interference as a function of the interviewer's race. These results support the conclusion that stereotype monitoring utilizes working memory capacity limiting the remaining resources available for other tasks.

Taken together these studies demonstrate that cultural knowledge of stereotypes and stereotype activation are ubiquitous. Furthermore, the influence of stereotypes in judgments and decision making is most pronounced when capacity is limited as stereotype use is energy efficient whereas stereotype avoidance requires cognitive resources.

1.12 Stereotypes Bias Memory and Decision Making

"The possession of stereotypes may interfere with even the simplest rational judgment."

(Allport, 1954)

Stereotypes can promote event-memory errors. For example, Kleider, Pezdek, Goldinger, and Kirk (2008), presented participants with a slideshow depicting a homemaker and a handyman working in a home performing both stereotypical (handyman hammering nails) and atypical (handyman baking cookies) actions. Two days later, participants were more likely to misattribute actions to the stereotypical actor as opposed to the atypical actor witnessed performing the action. The authors suggested that people rely on stereotypes to determine the source of information when episodic memory is weak.

Stereotypes can also promote face recognition errors. Oliver (1999) showed participants a 30-minute news clip wherein two wanted posters were displayed for 10 seconds each. The posters depicted Black and White perpetrators of one violent and one nonviolent crime (counterbalanced). The participants tried to identify the violent criminal among foils both immediately and again three months later. The accuracy of the immediate identifications was not affected by the race of the perpetrator. However, over time, participants that originally saw a White suspect paired with the violent act were increasingly likely erroneously to recall a black suspect. In a follow-up study, Oliver and Fonash (2002) decreased the time delay between witnessing the perpetrator and completing the identification task. Participants read crime briefs coupled with a photograph of the suspect (both Black and White). After a 20-minute delay, participants were asked to identify which of 10 photos were featured in the previously read crime briefs. Of the 10 presented photos only 4 were suspects (6 were filler, 3 Black, and 3 White). Participants were significantly more likely to misidentify a Black foil than a White foil as featured in the violent crime stories. The authors found no difference in identifications for the nonviolent stories.

Stereotypes can bias decision making. Jones and Kaplan (2003) found that when mock-jurors were asked to make verdict decisions for perpetrators accused of committing stereotype-consistent (i.e., Black man committing a blue-collar crime like grand-theft auto or a White man committing a white-collar crime like embezzlement) versus stereotype-inconsistent crimes participants rendered more guilty verdicts, gave longer sentences, held the perpetrator more responsible for the crime, and were more confident in their decisions. The authors theorized that stereotype-confirming information elicits less of a search for disconfirming facts.

These findings together highlight the point that stereotypes can have a biasing influence on various cognitive processes. Furthermore, these studies emphasize the importance of available cognitive resources in combating stereotypical biases. Thus, it seems highly plausible that individual WMC may well be a predictor of stereotypical influence. However, with the exception of one study (described below), extant literature has yet to establish such a link.

1.13 Individual Working Memory Capacity (WMC) and Stereotype Use

Kleider and colleagues (2012) presented high and low WMC participants, playing the role of mock-jurors, with ostensible trial case facts. Each case was shown concurrently with a photo of the accused defendant in the crime. Participants were asked to read the case facts and then render a verdict decision of guilt or innocence and indicate their level of confidence in their verdict decision (1-7). Decision time was unlimited. The accused defendant was either a Black or a White man. All cases depicted Black stereotype-consistent crimes (e.g., armed robbery). Additionally, for half of the cases, participants were required to hold an additional load in working memory (i.e., six to-be-remembered nonsense words) during the decision making portion of the task. Finally, individual prejudice level was assessed.

Kleider and colleagues (2012) suggested that while race should not be an influencing factor in juror-type judgments, knowledge of the Black-man-as-criminal stereotype is ubiquitous in American culture; thus, avoiding a biased decision requires cognitive resources. They hypothesized a high level of prejudice would act as an additional load as monitoring biases requires working memory capacity (Richeson & Shelton, 2003). Thus, they posited that low WMC individuals who were high in prejudice against Black persons would render stereotype-consistent (i.e., biased against Black defendants) verdict responses but only when cognitively taxed (i.e., externally loaded). The high WMC persons were not expected to make biased verdict decisions regardless of individual prejudices or external load manipulations.

The results supported these predictions. Under conditions of high external load, participants that were high in individual prejudice and low in WMC rendered significantly harsher judgments to Black compared to White defendants. The high WMC persons, regardless of individual prejudice level, were able to render similar judgments to the Black and the White defendants under conditions of external load. When sufficient cognitive resources were available (i.e., no external load conditions), high prejudice participants rendered equivalent verdicts to Black and White defendants regardless of individual WMC. These findings support the suggestion that high WMC is protective against the ubiquitous influence of stereotype in deliberative decision making.

1.14 The Current Study

In the current study I attempted to extend the findings of Kleider et al (2012) by testing whether persons with low relative to high WMC were more likely influenced by stereotypes when making judgments in maintenance and retrieval tasks. Specifically, Unsworth and Engle's (2007) maintenance and retrieval theory of individual differences in WMC stipulated conditions under which high WMC persons should outperform low WMC persons, including tasks requiring that 1) goal relevant information be maintained in primary memory, or 2) delimited controlled searches be conducted into secondary memory. Given that stereotypical associations are automatic and cognitively efficient, I posited that without sufficient resources to suppress, low relative to high WMC persons would be more likely to make more stereotype-consistent errors on maintenance and retrieval tasks. However, as explained above, low and high WMC persons were not expected to perform differently on other cognitive tasks (Unsworth & Engle, 2007). Thus, performance and stereotype use were predicted to be similar for low and high WMC persons completing other decision making tasks that required inhibiting a response and recognizing stimuli as familiar.

In the present investigation, the influence of stereotypes on cognitive tasks as a function of individual WMC was explored in a five-part study. First, in a Pilot Experiment I asked whether all participants (low and high WMC) have equivalent knowledge of common racial stereotypes. Then, to answer the primary questions of interest, I conducted four experiments. The first set of experiments involved cognitive tasks using primary memory; one maintenance and one inhibition task. With a single-task manipulation, I hypothesize that WMC affects performance and stereotype influence on maintenance tasks but not inhibition tasks. The second set of experiments involved the use of secondary memory; one retrieval and one familiarity task. Again, with a single-task manipulation, I aim to demonstrate that WMC affects performance and stereotype influence on retrieval tasks but not familiarity tasks. To that end, participants' WMC was tested using the operation span task (Unsworth, et al., 2005). Next, participants completed a cognitive maintenance (Experiment 1A) or inhibition task (Experiment 1B) and retrieval (Experiment 2A) or familiarity task (Experiment 2B). Finally, individual prejudice was measured.

CHAPTER 2:

PILOT EXPERIMENT

2.1 Overview

The aim of the Pilot Experiment was to determine whether low compared to high WMC persons maintain different knowledge regarding common racial stereotypes. To test this, high and low WMC participants were presented with a series of person-descriptors (e.g., athletic, disciplined) wherein they determined whether each descriptor reflected a racial stereotype of either Black persons, Asian persons, or neither. I predicted that persons with low and high WMC and low and high levels of individual prejudice would demonstrate high agreement regarding common racial stereotypes. That is, I suggest that descriptors considered highly reflective of the Black and/or Asian racial stereotype would be similar for all participants regardless of individual WMC or prejudices.

Participants completed a total of three tasks in the following order: Operation span task (Opspan), stereotype agreement task, and individual prejudice survey. The Opspan task required participants to solve a series of two-step math equations while holding a list of letters in memory. Next, participants determined whether they thought a series of descriptors accurately reflected a common societal stereotype by placing each descriptor in a labeled 'bucket' (i.e., Black, Asian, Neither). Finally, the Bogardus (1925) Social Distance Scale (SDS) was administered as a measure of individual prejudice.

2.2 Methods

2.2.1 Participants

A total of 119 (110 female)[3] undergraduate students, who self-identified as native English speakers, were recruited from introductory psychology courses at Georgia State University in exchange for course credit. Participants were at least 18 years of age. No other exclusion criterion was used. Participants were run in groups of 2 to 8 during experimental sessions.

2.2.2 Materials

2.2.2.1 Operation Span Task Participants first completed the Automated Operation Span Task (OPSPAN; Unsworth, et al., 2005). In this computer-administered test that assesses WMC, an equation (e.g., “9/3 + 5 = 8”) is presented and participants are prompted for a true-false response. Next, a to-be-remembered letter is displayed for 800ms. The equation/letter sequence is repeated until 2-7 letters have been presented. When prompted, participants are required to recall all letters, in presentation order, via mouse-click. On-screen feedback displays the total number of correctly recalled letters for each trial. There are a total of 75 equation/letter trials that vary in set size. The OPSPAN score is the number of correctly recalled letter sequences across all trials. Higher scores reflect higher WMC.

2.2.2.2 Stereotype Agreement Task Participants were presented with a series of 100 descriptors on a computer screen. Included descriptors were expected to be considered common stereotypes for African American (41) and Asian men (39) as well as neutral (20). Descriptors expected to be stereotype-consistent were chosen based on pre-pilot questionnaires and previous research

---

[3] Although females were disproportionate represented in the current sample, it is important to note this this sample is demographically similar to the other samples included in this study.

(Stewart, Weeks, and Lupfer, 2003; see Table 1 for complete list). Participants were asked to indicate whether each descriptor reflected an existing racial stereotype for African American men, Asian men, or neither. Participants indicated their decisions by pressing the labeled key on the keyboard.

2.2.2.3 Individual Prejudice Survey The Bogardus (1925) Social Distance Scale (SDS), an established self-report measure of prejudice attitudes towards African Americans, was used to asses individual prejudice levels. The scale was modified in order to also obtain self-report measures of prejudice towards Asian Americans. Participants were presented with the phrase 'I would be willing to have a White American person as my:' followed by a sequence of 14 nouns (e.g., next door neighbor, romantic date, governor, wife or husband). Participants rated the degree to which they agreed with each statement using a 9 point Likert-type scale. Next, participants were presented with the phrase 'I would be willing to have a Black American person as my:' followed by the same sequence of 14 nouns. Finally, participants were presented with the phrase 'I would be willing to have an Asian American as my:' again followed by the same sequence of 14 nouns. Statement presentation and response time was unlimited.

2.2.3 Procedure

Participants signed consent forms and were seated at one of eight individual computer stations. Participants were informed that the study consists of several tasks and that the specific directions for each task would be read as the session progressed.

2.2.3.1 Operation Span Task Participants were told that for the first task we were interested in their ability to do two things at once. Practice trials were presented in order to ensure that participants were comfortable with the task speed. During the test trials, their job was to solve the math equations quickly and accurately while holding the letter lists in memory until prompted to recall the letters in presentation order. All responses were made via mouse-click.

2.2.3.2 Stereotype Agreement Task Participants were informed that for the next task we were interested in what people consider to be common racial stereotypes. Specific instructions read as follows, "The next thing you are going to be doing is reviewing a series of descriptors sometimes ascribed to specific ethnic groups. Some are commonly known stereotypes and others are not. Please determine if each descriptor reflects a common social stereotype, and if so, then for what group. While making these decisions, please keep in mind that we are not asking whether or not you agree with the implied association but rather, if you agree that is in fact a known stereotype." Participants indicated their response by pressing the labeled key on the keyboard. Response time was unlimited. Once a response was logged, the next descriptor was presented.

2.2.3.3 Individual Prejudice Survey Finally, the Bogardus (1925) Social Distance Scale (SDS) was administered. Participants were told that they were answering questions regarding their opinions on various social issues. The true purpose of the measure was concealed until debrief in an effort to avoid biasing participant responses through the influence of demand characteristics or social desirability concerns. The session concluded with a brief demographic survey.

Following all tasks, participants were debriefed, thanked, and dismissed.

2.3 Results

2.3.1 OPSPAN Scores

A tertiary split was used to separate participants into WMC categories, following the method of Goldinger, et al., (2003). Approximately one-third of the original participants were considered to have low WMC (n= 41; WMC scores ranging from 6-24) and one-third were considered to have high WMC (n= 38; WMC scores ranging from 50-75), the remaining one-third (n= 40) were considered to have average WMC and were not included in relevant comparisons.

2.3.2 Prejudice Levels

Two prejudice scores were calculated for each participant; one (high, low) for prejudice towards Black Americans and one (high, low) for prejudice towards Asian Americans. Prejudice levels were calculated by summing participant responses to each of 14 statements (i.e., 1-9) separately for the three question-types (White American persons, Black American persons, and Asian American persons). The response total of the questions pertaining to Black American persons was subtracted from the response total of the questions pertaining to White American persons. This calculation was performed separately for each participant. Negative and zero prejudice scores indicate a bias in favor of Black persons or a neutral attitude and were considered to reflect low levels of individual prejudice (n= 85) whereas positive prejudice scores indicate a bias against Black persons and were considered to reflect high levels of individual prejudice (n= 34). Next, the response total of the questions pertaining to Asian American persons was subtracted from the response total of the questions pertaining to White American persons. Again, the calculation was performed separately for each participant. Negative and zero prejudice scores indicate a bias in favor of Asian persons or a neutral attitude and were considered to reflect low levels of individual prejudice (n= 46) whereas positive prejudice scores indicate a bias against Asian persons and was considered to reflect high levels of individual prejudice (n= 73). [4]

2.3.3 Stereotype Knowledge Survey

I had multiple goals for the current dataset. First, I wanted to identify the list of descriptors that were considered stereotype-consistent for African American men and Asian American men as well as those that were considered neutral. To that end, the percentage of participants identifying a descriptor as a common social stereotype for each group (African American, Asian, and neither) was calculated. A descriptor was considered to be stereotype-consistent or neutral if at least 50% of participants identified it as common (or not) to a specific group. The entire list of presented descriptors is displayed in Table 1, as well as how they were originally predicted to be identified, and which predictions were supported. Overall, 72% of the descriptors were identified consistent with expectations; particularly for the descriptors posited to be common African American (95%) and Asian stereotypes (79%). However, the descriptors predicted to be neutral were rarely identified as such (20%). Thus, we generated a new set of descriptors expected to be neutral and tested them on a new set of participants (discussed in detail below). The complete list of stereotype-consistent descriptors and agreement ratings are available in Table 2.

My second goal was to identify descriptors that could be used in memory portions of later experiments. Thus, I had two criteria when considering which descriptors to include. First, the majority of persons had to agree that the descriptor reflected a common social stereotype, and

[4] In all analyses for Experiments 2 and 3, prejudice was controlled for; however, in the current analyses it was a between-subjects factor in order to ensure that individual prejudice level did not impact knowledge of common social stereotypes.

second, given that many of the descriptors were synonyms, I could only have a limited number of descriptors with overlapping meanings. The final set of descriptors used for the memory tests are available in Table 3.

My third goal was to test my prediction that high and low WMC persons have similar knowledge of common racial stereotypes. To that end, I looked at the percentage of high, compared to low, WMC persons that acknowledged common racial stereotypes. My focus for these comparisons were the descriptors identified as to-be-used in Experiment 2 as these are the descriptors we want to ensure persons maintain equivalent stereotypical knowledge for regardless of WMC. Knowledge was considered to be equivalent across groups so long as at least 51% of low and high WMC participants identified it as common to the group identified as consistent (e.g., Football Player and African American). All but three descriptors passed this test (meticulous: low WMC- 44%, high WMC- 61%; mild: low WMC- 49%, high WMC- 53%; submissive: low WMC- 49%, high WMC- 58%).

2.3.4 New Neutral Descriptors

As the majority of the original 'neutral' descriptors (80%) were not classified as belonging to the 'neither' category as expected, we generated a new list of descriptors/phrases that we tested for their neutrality. Twenty-three participants (a subset of those included in the previous rating task) rated 135 words/phrases (100 used in original rating task plus 35 new neutral words/phrases) using the identical protocol described above. The new neutral list can be seen in Table 4. Overall, 58% of the new neutral words were not considered to be common racial stereotypes for either African American men or Asian men and were therefore considered neutral. Eighteen of the new neutral words were chosen to be included in later memory tasks.

2.4 Discussion

Overall, these findings suggest that there is generally high agreement regarding common racial stereotypes among both high and low WMC persons. In the current study, high and low WMC persons were considered to be in agreement that a stereotype was common to a specific ethnic group so long as 51% or more of each group agreed. However, it is unknown if and/or how the differences in degree of agreement may have affected the results of the memory test (Experiments 3a and 3b). That is, although the majority of all persons, regardless of WMC, were in agreement with regards to which descriptors were common to African American or Asian American men, sometimes more low, than high, and sometimes more high, than low, WMC persons indicated agreement e.g., 76% versus 81%). Although this is worth future study, I suggest that these differences do not preclude drawing valuable insights form Experiment 3 results.

CHAPTER 3:

EXPERIMENT 1 (A and B)

3.1 Overview

The goal for Experiment 1 (A and B) was to assess whether low WMC persons make more stereotype-consistent errors during a cognitive maintenance task than high WMC persons. To that end, participants completed a modified go/no-go task (Redick, et al., 2011). Specifically, participants were asked to respond 'go' to pictures that follow a rule set and inhibit responses to pictures outside the rule set. A series of pictures depicting faces (Black, White), objects (tool, weapon), and shapes (square, circle) were presented sequentially. A rule set determined which face-type, object-type, and shape-type warranted a go-response. The rule set included either stereotype-consistent 'go-targets' (Black face and weapon) or stereotype-inconsistent 'go-targets' (Black face and tool). The shape category is considered to be neutral (neither stereotype-consistent nor stereotype-inconsistent). The purpose of the shape being included in the 'go' rule set is to serve as a baseline measure for individual maintenance errors. Accuracy and response time of go/no-go responses were recorded.

In Experiment 1 version A (Maintenance), the identity of an accurate go-stimulus (face-type, object-type, or shape-type) was contingent upon the identity of the previous go-stimulus (must come from a different 'go' category). Thus, withholding a go-response to lures (picture from the same go-category as the previous go-category) presents a need to override an automatic response. This automatic-response bias is expected to be further exacerbated by the stereotype-consistent lures (compared to stereotype-inconsistent or stereotype-neutral lures). Therefore, I suggest that fewer cognitive resources will lead to 1) increased errors (go-response to lures) overall and, 2) increased errors to stereotype-consistent lures specifically.

In Experiment 1 version B (Inhibition), as the task rules change, successful performance is dependent upon one's ability to inhibit a response. As performance on inhibition tasks have not been found to differ as a function of WMC, I expected that low and high WMC persons would perform equally well in the go/no-go responses. I predicted no performance differences in number of errors committed or influence of stereotypes on errors made as a function of individual WMC.

3.2 Method

3.2.1 Participants

A total of 248 (122; 106 female completed the maintenance task (A) and 126; 97 female completed the inhibition task (B)) undergraduate students were recruited from introductory psychology courses at Georgia State University in exchange for course credit. Participants were at least 18 years of age. No other exclusion criterion was used. Participants were run in groups of 2 to 8 during experimental sessions.

3.2.2 Materials

3.2.2.1 Operation Span Task Participants first completed the Automated Operation Span Task (OPSPAN; Unsworth, et al., 2005) described above.

3.2.2.2 Maintenance Task (A) The maintenance task used in the present investigation was modeled after the conditional go/no-go task used by Redick and colleagues (2011) wherein low WMC persons were found to commit more errors than high WMC persons. Specifically, individual picture stimuli consisting of faces (Black and White), objects (weapons and tools), and shapes (circles and squares) were presented in the center of a computer screen. Participants were instructed to respond to go-stimuli (those following a given rule set) by using their dominant hand to press a single button on the keyboard. A rule set established one face-type (Black or White), one object-type (weapon or tool), and one shape-type (circle or square) that warranted a go-response. Furthermore, the rule specified that only pictures from a category different than the category of the previous go-trial require a go-response. All pictures outside the stated rule set require a no-go response (see Figure 1 for pictorial explanation). Responses to lure presentations are of primary interest. Lures consisted of any pictures included in the go- rule set but which require a no-go response. For example, if a participant's first go-response is to a go-object then their next go-response should be to a go-face or go-shape; however, if the next go-target presented is a go-object (i.e., a lure), a no-go response is the correct choice. The rule set included either stereotype-consistent targets (e.g., Black face and weapon) or stereotype-inconsistent targets (e.g., Black face and tool). The stereotype-consistency of the rule set varied between subjects. All rule sets include a neutral target (e.g., circle). Each picture was shown in color in the center of a black screen for 450ms, followed by a by a blank screen for 700ms. Participants had 1150ms to make a response to each picture.

After completing 40 practice trials, participants completed three test blocks, each consisted of 200 trials. The trials included 40% distractor trials (any pictures not included in the 'go' rule set), 40% target trials, and 20% lure trials (10% stereotype-consistent, 10% stereotype-neutral). Each block took approximately 4 minutes to complete.

3.2.2.3 Inhibition Task (B) The inhibition task used in the present investigation was modeled after the go/no-go task used by Redick and colleagues (2011) wherein low and high WMC persons were found to commit a similar number of errors. The task was virtually identical to the maintenance task with a single modification; a go-response was not contingent upon the category membership of the previous go-response. The rule set again established one face-type (Black or White), one object-type (weapon or tool), and one shape-type (circle or square) that warranted a go-response. All pictures outside the stated rule set required a no-go response. The rule set either included stereotype-consistent targets (e.g., Black face and weapon) or stereotype-inconsistent targets (e.g., Black face and tool). The rule set changed half-way through the task such that persons originally responding 'go' to Black faces and weapon objects, next responded 'go' to Black faces and tool objects, and vice versa. The presentation order of the rule set varied between subjects. All rule sets included a neutral target (e.g., circle). Each picture was shown in color in the center of a black screen for 450ms, followed by a by a blank screen for 700ms. Participants had 1150ms to make a response to each picture.

After completing 40 practice trials, participants completed the first 200 trial test block, next the rule set was changed and participants again completed 40 practice trials and finally the second 200 trial test block. The trials included 20% distractor trials (any pictures not included in the 'go' rule set), 80% target trials.

3.2.2.4 Individual Prejudice Survey The Bogardus (1925) Social Distance Scale (SDS), described above, was also used.

3.2.3 Procedure

Participants signed consent forms and were seated at one of eight individual computer stations. Participants were informed that the present study consists of several tasks and that the specific directions for each task would be read as the session progressed.

3.2.3.1 Operation Span Task Participants were told that for the first task we were interested in their ability to do two things at once. Practice trials were presented in order to ensure that participants were comfortable with the task speed. During the test trials, their job was to solve the math equations quickly and accurately while holding the letter lists in memory until prompted to recall the letters in presentation order. All responses were made via mouse-click.

3.2.3.2 Maintenance Task (A) Participants were informed that the next task was a go/no-go task wherein they would be presented with a rule set and a series of photos. The rule set was explained as follows: “the picture series will include photos of Black and White faces, circle and square shapes, and weapon and tool objects. Your job is to respond ‘go’ only to Black faces, circle shapes, and weapon (or tool) objects (go-targets will be varied across participants). Furthermore, you will only respond ‘go’ to a go-target (i.e., Black face, circle shape, weapon object) that comes from a category (i.e., face, shape, or object) different than the previous go-target. For example, if you see a circle shape and respond ‘go’ then the next time that you respond ‘go’ should be to either a Black face or a weapon object but NOT to a circle shape. Thus, targets that require a ‘go’ response depend on the category of the previous go-target. Do these instructions make sense? Do you have any questions? ” Participants were instructed to respond ‘go’ to the photos that follow the rule set by pressing a marked key on the keyboard with their dominant hand. They were to withhold a response to any photo outside the rule set. Participants were informed that they were to make their responses as quickly and accurately as possible. The specific rule set guiding their responses as well as task instructions were reiterated on the computer screen prior to the start of the task. Practice trials were presented in order to ensure that participants were comfortable with the task speed. Once all participants completed the practice block, the experimenter offered anyone that was not comfortable moving onto the test trials the opportunity to repeat the practice block before moving forward. Response accuracy and reaction time was recorded.

3.2.3.3 Inhibition Task (B) Participants were informed that the next task was a go/no-go task wherein they would be presented with a rule set and a series of photos. The rule set was explained as follows: “the picture series will include photos of Black and White faces, circle and square shapes, and weapon and tool objects. Your job is to respond ‘go’ only to Black faces, circle shapes, and weapon (or tool) objects (go-targets will be varied across participants). Do these instructions make sense? Do you have any questions? ” Participants were instructed to respond ‘go’ to the photos that followed the rule set by pressing a marked key on the keyboard with their dominant hand. They were to withhold a response to any photo outside the rule set. Participants were informed that they should make their responses as quickly and accurately as possible. The specific rule set guiding their responses as well as task instructions was reiterated on the computer screen prior to the start of the task. Practice trials were presented in order to ensure that participants were comfortable with the task speed. Once all participants confirmed that they were comfortable with the task, they moved forward with the test trials. Response accuracy and reaction time was recorded.

3.2.3.4 Individual Prejudice Survey Finally, the Bogardus (1925) Social Distance Scale (SDS) was administered. Participants were told that they would be answering questions regarding their opinions on various social issues. The true purpose of the measure was concealed until debrief in an effort to avoid biasing participant responses through the influence of demand characteristics or social desirability concerns. The session concluded with a brief demographic survey.

Following all tasks, participants were debriefed, thanked, and dismissed.

3.3 Results

3.3.1 OPSPAN Scores

A tertiary split was used to separate participants into WMC categories, following the procedure of Goldinger, et al., (2003). Approximately one-third of the original participants were considered to have low WMC (n = 42 maintenance task, range: 0-32; n = 41 inhibition task, range: 0-33) and one-third were considered to have high WMC (n = 39 maintenance task, range: 51-75; n = 43 inhibition task; range: 47-75), the remaining one-third (n = 41 maintenance task; n = 42 inhibition task) were considered to have average WMC and were not included in the analyses.

3.3.2 Prejudice Levels

Prejudice levels were calculated by summing participant responses to each of 14 statements (i.e., 1-9) separately for the two question-types (White American persons and Black American persons). The response total of the questions pertaining to Black American persons was subtracted from the response total of the questions pertaining to White American persons. This calculation was performed separately for each participant. Prejudice score was controlled for in all analyses.

3.3.3 Maintenance Task (A)

Two 2 Rule Set (stereotype-consistent, -inconsistent) x 2 WMC (high, low) ANCOVAs were conducted with 1) target hits and, 2) distractor errors as the dependent variables, controlling for individual prejudice score. The results revealed a significant rule set x WMC interaction on target hits $F(1,76) = 6.65$, $\eta_p^2 = .08$, $p = .01$, as low WMC persons made more target hits in the stereotype-consistent (88%) than –inconsistent (78%) rule condition $F(1,39) = 5.36$, $\eta_p^2 = .12$, $p = .03$. Additionally, low WMC persons made fewer target hits in the stereotype-inconsistent condition (78%) than high WMC persons (87%) $F(1,39) = 4.5$, $\eta_p^2 = .10$, $p = .04$; whereas high WMC persons attained similar hits across conditions (stereotype-consistent: 82%, stereotype-inconsistent: 87%) and hit rate was statistically equivalent within the stereotype-consistent condition (low WMC: 88%, high WMC: 82%). There were no significant effects or interactions for error rate to distractors.

Next, the overall 2 Lure-Type (stereotype-related, stereotype-neutral) x 2 Rule Set (stereotype-consistent, -inconsistent) x 2 WMC (high, low) ANCOVA was conducted, controlling for individual prejudice score. Error rate to lures were the dependent variable of interest. There was a main effect of lure-type, with persons committing significantly more errors to stereotype-related (M = 50%) than –neutral lures (M = 46%), $F(1,76) = 10.99$, $\eta_p^2 = .13$, $p < .01$. The main effect of WMC was significant $F(1,76) = 5.51$, $\eta_p^2 = .07$, $p = .02$, as low WMC persons committed more errors (54%) than high WMC persons (42%). Finally, the lure-type x WMC interaction was also significant $F(1,76) = 4.23$, $\eta_p^2 = .05$, $p = .04$, as low, compared to high, WMC persons committed a greater number of errors to stereo-related lures (low WMC: 57%, high WMC: 43%) while all persons, regardless of WMC, committed a similar number of errors to the neutral lures (low WMC: 50%, high WMC: 41%).

Next, I tested my three a-priori predictions. First, low WMC persons should commit more errors to stereotype-consistent than –inconsistent lures. Results failed to support this prediction as low WMC persons committed a similar number of errors to stereotype-consistent (M = 61%) than –inconsistent (M = 54%) lures $F(1,39) = 1.03$, p = ns. Second, low WMC persons should commit more errors to stereotype-consistent than neutral lures. Support for this prediction was generated as low WMC persons committed more errors to stereotype-consistent (*M* = 61%) than neutral (*M* = 55%) lures $F(1,19) = 5.19$, $\eta_p^2 = .22$, $p = .03$. Finally, low WMC persons should commit more errors to stereotype-consistent lure than high WMC persons. Supporting this prediction, low, compared to high, WMC persons committed more errors to stereotype-consistent lures (low WMC: 61%, high WMC: 40%) $F(1,36) = 9.8$, $\eta_p^2 = .21$, $p < .01$. No other simple effect comparisons were significant (see Figures 2 and 3 for a pictorial representation of findings). Together, these results provide strong support for the suggestion that low WMC persons are more susceptible to making stereotypical errors on maintenance tasks than either neutral errors or high WMC persons.

3.3.4 Inhibition Task (B)

I expected that low and high WMC persons would perform equally well on both target and distractor trials regardless of whether the rule set included stereotype-consistent or –inconsistent targets. To test this prediction, two 2 Picture-Type (stereo-related, neutral) x 2 Rule Set (stereotype-consistent, -inconsistent)[5] x 2 WMC (high, low) ANCOVAs were conducted, controlling for individual prejudice score. Target hits and distractor errors were the dependent variables of interest. Picture-Type served as the within-subjects factor while rule set and WMC were between-subjects factors. When considering target hits, there were no significant main effects or interactions. Regarding distractor errors, a main effect of picture-type emerged, as persons were more likely mistakenly to respond ‘go’ to a stereo-related (35%) than a neutral (27%) picture, $F(1,79) = 14.32$, $\eta_p^2 = .15$, $p < .01$. There were no other significant main effects or

[5] Only performance on the second set was considered as this is where inhibition was required to overcome the previously learned and practiced rule.

interactions.[6] Overall, these results support the hypothesis that WMC does not affect performance differences on inhibition tasks regardless of stereotype-consistency.

3.4 Discussion

Together, the significance of the abovementioned results are twofold. First, they provide strong support for my suggestion that low WMC persons are more susceptible to stereotypical errors than high WMC persons when performing maintenance tasks. Second, they support and extend the maintenance portion of Unsworth and Engle's (2007) maintenance and retrieval theory of individual differences in WMC, as WMC was found to influence errors made when performing maintenance but not inhibition tasks. However, it was surprising that low WMC persons committed a similar number of errors to stereotype-consistent and –inconsistent lures. Potentially, the stereotype-consistent compared to the –inconsistent rule set promoted an increased 'go' response leading to a greater number of hits to 'go' targets and false alarms to 'no-go' lures. That is, maybe the stereotype-consistent targets naturally fit together better than the stereotype-inconsistent targets causing low WMC persons, who maintain insufficient resources to override an automatic response, to adopt a more liberal 'go' criterion in the stereotype-consistent condition (Signal Detection C; see Macmillan and Creelman, 2005).

To test this suggestion, hit rate (correctly saying 'go' to target items) and false alarm rate (incorrectly saying 'go' to lure items) were calculated, and then analyzed with signal detection procedures to ascertain accuracy (*d')* and bias (C) scores. Higher *d'* scores indicate better

---

[6] The Picture-type x rule set interaction was trending toward significance $F(1,79) = 3.45$, $\eta_p^2 = .04$, $p = .07$ with persons following stereotype-consistent rules committing 38% errors to stereo-related pictures and 27% errors to neutral pictures and persons following stereotype-inconsistent rules committing 32% errors to stereo-related pictures and 27% errors to neutral pictures. There was not a main effect of rule set for either stereo-related or neutral pictures.

sensitivity than lower scores. *C* scores represent the criterion, or response bias, used to identify a target. The criterion scores range from conservative/strict (1.0), suggesting a tendency to not say 'go' to an item, to liberal/lax (-1.0), suggesting a tendency to say 'go' to an item. A 2 Rule Set (stereotype-consistent, -inconsistent) x 2 WMC (high, low) between-subjects ANCOVA was conducted with d' and C as the outcome variables of interest. There was a main effect of WMC on d' $F(1,76) = 5.94$, $\eta_p^2 = .07$, $p = .02$, as high WMC persons were better able to distinguish between hits and false alarms (1.29) than low WMC persons (.91). Consistent with post-hoc predictions, there was a significant two-way interaction for differences in C, $F(1,76) = 4.90$, $\eta_p^2 = .06$, $p = .03$, such that in the stereotype-consistent rule set low versus high WMC persons adopted a significantly more liberal 'go' criterion (low WMC: -.77, high WMC: -.36). Additionally, the tendency for low WMC persons to adopt a more liberal decision criterion when in the stereotype-consistent ($M = -.77$), versus –inconsistent ($M = -.46$), rule condition was trending toward significance ($p = .08$). Conversely, high WMC persons adopted a similar decision making criterion regardless of rule set (stereotype-consistent: -.36, stereotype-inconsistent: -.52). Thus, although low WMC persons committed a similar number of errors to lures regardless of rule set, this result suggests a trend for the stereotype-consistent rule set to facilitate a 'go' response whereas the stereotype-inconsistent rule set inhibited a 'go' response.

CHAPTER 4:

EXPERIMENT 2 (A and B)

4.1 Overview

The goal for Experiment 2 (A and B) was to assess whether low WMC persons make more stereotype-consistent errors during a memory retrieval task than high WMC persons. To that end, participants were presented with multiple to-be-remembered word lists. Each word list consisted of ten phrases describing an ostensible person. The person's name was at the top of the list to act as a category label. Two of the names were considered highly representative of African American male names (e.g., DeShawn Washington) and two of the names were considered highly representative of Asian male names (e.g., Li Nguyen). The remaining two names were control names (e.g., David Williams) and not associated with or indicative of a particular race. The ten phrases included stereotype-consistent descriptors, stereotype-inconsistent descriptors, and stereotype-neutral descriptors. Phrase type was determined through pilot ratings (see Pilot Experiment above). Memory for the phrase descriptors was the dependent variable of interest.

Similar to Experiment 1, versions A and B of the experiment were identical with the exception of a single task manipulation. Following the guidelines outlined in Unsworth and Engle's (2007) framework, procedures used in version A were designed to require a delimited search into short-term memory (i.e., active retrieval), and thus, should bring about performance differences between low and high WMC persons. Specifically, following each list presentation, participants were asked to complete a free recall memory test. Conversely, version B required that participants recognize presented descriptors as 'old' and previously presented or 'new' and never before seen. Familiarity promotes successful task completion; thus, low and high WMC persons were expected to perform equivalently.

In Experiment 2 version A (Retrieval), accurate retrieval of to-be-remembered items across multiple lists including both stereotype-consistent and -inconsistent targets required a controlled search using delimited cues (descriptors only from the most recently presented list), thus, I posited that less specific search cues (characteristically used by low WMC persons) would lead to poorer performance on memory tests. Therefore, I expected that when recalling list phrases, regardless of individual prejudice level, low compared to high WMC persons would demonstrate a greater number of stereotype-consistent intrusions from previously studied lists and/or a lesser number of stereotype-inconsistent hits (i.e., accurate recall) from presently studied lists.

In Experiment 2 version B (Familiarity), successful performance relied on one's sense of familiarity. As judging familiarity is an automatic response, and not dependent upon cognitive resources (Oberauer, 2005), I expected that low and high WMC persons would be equally reliable in determining whether a presented word was 'old' and previously shown or 'new' and never before seen. I predicted no performance differences in number of errors committed or in the influence of stereotypes on errors made as a function of individual WMC.

4.2 Method

4.2.1 Participants

A total of 214 (112; 86 female completed the retrieval task (A) and 102; 88 female completed the familiarity task (B)) undergraduate students, who self-identified as native English speakers, were recruited from introductory psychology courses at Georgia State University in exchange for course credit. Participants were least 18 years of age. No other exclusion criterion was used. Participants were run in groups of 2 to 8 during experimental sessions.

4.2.2 Materials

4.2.2.1 Operation Span Task Participants first completed the Automated Operation Span Task (OPSPAN; Unsworth, et al., 2005) described above.

4.2.2.2 Retrieval Task (A) The retrieval task used in the present investigation was modeled after the task used by Kane and Engle (2000) wherein low WMC persons were found to commit more recall errors than high WMC persons. The present retrieval task included six word list presentations. For each list presentation participants were shown 10 phrases/descriptors describing an ostensible target person. After the list presentation, participants completed a rehearsal prevention task before being asked to recall and record as many of the descriptors as possible.

Each list was included 10 descriptors intended to describe an ostensible person (2 Black, 2 Asian, 2 ambiguous). The ethnicity of the person was implied through the name (first name and surname) which was presented akin to a category label for each word list. The ethnic surnames were chosen by looking at demographic data online. Surnames that were identified as one of the 200 most popular in America and most common among a specific ethnic group were chosen. For example, Washington was considered a popular surname and 90% of persons with the surname Washington self-identified as Black. The first names were taken from online lists detailing 'most common names for Black/Asian boys.' The ambiguous names also included surnames that were among the 200 most popular in America but no ethnicity claimed the majority of these names. For example, Williams was considered a popular surname, however, 49% of persons with the surname Williams self-identified as White, 47% Black, >1% Asian, and >2% Hispanic. Additionally, the ambiguous first names were considered popular boy names but did not fall on either the 'Black' or 'Asian' list. Pilot testing was used in order to confirm that the names chosen for each of the three ethnicity categories (Black, Asian, ambiguous) were appropriate. Eight judges determined that DeShawn Washington and Jamal Jackson were Black (88% and 100% respectively), and Li Nguyen and Yen Tran were Asian (88% and 100% respectively). However, judges primarily thought that David Williams and John Woods were White (75% and 88% respectively).

Each list describing a racially-distinguishable target included 10 descriptors; 4 stereotype-consistent, 4 stereotype-inconsistent, and 2 neutral. Lists describing a racially-ambiguous target also included 10 descriptors: 4 stereotype-consistent to African American men, 4 stereotype-consistent to Asian men, and 2 neutral. The descriptor-type was determined based on pilot testing. All lists were presented on a computer screen with the person's name at the top of the screen and the descriptors shown one at a time in the center of the screen. Each descriptor was presented for 1750ms followed by a 250ms black screen. List presentation was complete after approximately 20 seconds.

Immediately after each list presentation participants completed a rehearsal prevention task for 16 seconds. The Trail-Making task (used by Kane and Engle, 2000; created independently by Ricker and Axelrod, 1994; Baddeley, 1996) presents participants with a letter and number target (e.g., F-61) and requires that they count up (e.g., G-62, H-63, I-64, etc.) as quickly and accurately as possible. After 16 seconds, participants were cued to type their ending results.

Next participants were allotted 30 seconds to retrieve the list words. For the memory test, participants were cued, after the rehearsal prevention task, to retrieve and list as many descriptors as possible from only the most recent list presentation. Participants recorded their retrieved word lists on a screen that listed only the targets first and last name. At the end of the 30 second retrieval period, the computer screen changed to indicate the start of the next trial.

After all lists were presented and free recall was been tested, participants completed a cued recall test. During the cued recall test each of the 60 descriptors listed during the presentation phase were show individually on the computer screen. With each descriptor, participants were prompted to indicate, via key-press, which individual the word described. The names of all six individuals were listed (alphabetically by surname) on the screen concurrently with each descriptor.

4.2.2.3 Familiarity Task (B) This task was identical to the retrieval task with only a minor change to the memory test; participant's recognition memory was tested. Specifically, following the presentation of a list of descriptors and the rehearsal prevention task, participants were shown a series of 6 (3 'old' and 3 'new') descriptors individually and asked to identify each as either 'new' and never before seen or 'old' and previously presented.

4.2.2.4 Individual Prejudice Survey The Bogardus (1925) Social Distance Scale (SDS), an established self-report measure of prejudice attitudes towards African Americans, was used to asses individual prejudice levels. We modified the scale in order to also obtain self-report measures of prejudice towards Asian Americans. Using E-Prime software participants were presented with the phrase 'I would be willing to have a White American person as my:' followed by a sequence of 14 nouns (e.g., next door neighbor, romantic date, governor, wife or husband). Participants rated the degree to which they agree with each statement using a 9 point Likert-type scale. Next, participants were presented with the phrase 'I would be willing to have a Black American person as my:' followed by the same sequence of 14 nouns. Finally, participants were presented with the phrase 'I would be willing to have an Asian American as my:' again followed by the same sequence of 14 nouns. Statement presentation and response time was unlimited.

4.2.3 Procedure

Participants signed consent forms and were seated at one of eight individual computer stations. Participants were informed that the present study consisted of several tasks and that the specific directions for each task would be read as the session progressed.

4.2.3.1 Operation Span Task For the first task, participants were told we were interested in their ability to do two things at once. Practice trials were presented in order to ensure that participants were comfortable with the task speed. During the test trials, their job was to solve the math equations quickly and accurately while holding the letter lists in memory until prompted to recall the letters in presentation order. All responses were be made via mouse-click.

4.2.3.2 Retrieval Task (A) Participants were informed that the next task was a multipart memory task. Participants were presented with a 10-phrase list connected to a target individual. List presentation took approximately 20 seconds (2 seconds per word). Following each list, participants completed a rehearsal prevention task (Trail-Making task) for 16 seconds. At the end of the Trail-Making task, participants were cued to record their ending result. Next the screen prompted participants to recall the word list with the following statement: You previously saw a list of words/phrases describing (DeShawn Washington). List as many of those words/phrases as you can recall on the next screen. You will be given 30 seconds to do this before the screen will automatically advance. Press the enter key after each word/phrase entry. After 7500ms the screen automatically advanced to a screen that prompted: (DeShawn Washington)… Participants were allotted 30 seconds to list as many words as possible after which the computer screen changed. Before the next trial began, a break screen was presented that stated: Please press the spacebar when you are ready to begin the next trial. The process repeated itself until all six word lists were presented and recalled. With the exception of the first trial, which always lists an ambiguous target person (and will serve as a baseline level of recall), the presentation order for the other lists was randomized.

Immediately following the free recall phase (including the presentation and free recall of six separate lists), participants completed a cued recall task wherein they were asked to indicate which of the six target persons each of the 60 presented adjectives described. Response time was unlimited.

4.2.3.3 Familiarity Task (B) Participants were informed that the next task was a multipart memory task. Participants were presented with a 10-phrase list connected to a target individual. List presentation took approximately 20 seconds (2 seconds per word). Following each list, participants completed a rehearsal prevention task (Trail-Making task) for 16 seconds. At the end of the Trail-Making task, participants were cued to record their ending result. Next, participants were prompted to identify a series of descriptors as either ‘new’ and never before seen or ‘old’ and previously presented. Responses were made via key press. The process repeated itself until all six word lists were presented and recognition memory was tested. With the exception of the first trial, which always lists an ambiguous target person, the presentation order for the other lists was randomized.

4.2.3.4 Individual Prejudice Survey Finally, the Bogardus (1925) Social Distance Scale (SDS) was administered. Participants were told that they would be answering questions regarding their opinions on various social issues. The true purpose of the measure was concealed until debrief in an effort to avoid biasing participant responses through the influence of demand characteristics or social desirability concerns. The session concluded with a brief demographic survey.

Following all tasks, participants were debriefed, thanked, and dismissed.

4.3 Results

4.3.1 OPSPAN Scores

A tertiary split was used to separate participants into WMC categories (Goldinger, Kleider, Azuma, and Beike, 2003). Approximately one-third of the original participants were considered to have low WMC (n = 35 retrieval task, range: 0-33; n = 35 familiarity task, range: 0-32;) and one-third were considered to have high WMC (n = 42 retrieval task, range: 46-75; n = 31 familiarity task, range: 51-75;), the remaining one-third (n = 35 retrieval task; n = 36 familiarity task) were considered to have average WMC and were not included in the analyses.

4.3.2 Prejudice Levels

Two prejudice scores were calculated for each participant; one for prejudice towards Black Americans and one for prejudice towards Asian Americans. Prejudice levels were calculated by summing participant responses to each of the 14 statements (i.e., 1-9) separately for the three question-types (White American persons, Black American persons, and Asian American persons). The response total of the questions pertaining to Black American persons was subtracted from the response total of the questions pertaining to White American persons. This calculation was performed separately for each participant. Next, the response total of the questions pertaining to Asian American persons was subtracted from the response total of the questions pertaining to White American persons. Again, the calculation was performed separately for each participant. Prejudice score was controlled for in all analyses.

4.3.3 Retrieval Task (A1) Low WMC persons were expected to make more recall errors (fewer hits and more intrusions) than high WMC persons. More specifically, low compared to high WMC persons were expected to demonstrate a greater number of stereotype-consistent intrusions from previously studied lists and/or a lesser number of stereotype-inconsistent hits. To test this prediction a mixed-model ANCOVA was run on the overall design: 2 Target-Race x 2 Word-Type x 2 WMC. Target-race (Black, Asian)[7] and word-type (stereotype-consistent, -inconsistent)[8] served as the within-subject factors whereas WMC (high, low) was between-subjects. Dependent variables of interest include accuracy calculated as a percent (accurate recall/total possible) and an intrusion rate count. Black American and Asian American prejudice levels were entered as control variables. See Tables 5 and 6 for overall recall accuracy/errors for all persons and all word types. Regarding memory accuracy, results revealed a main effect of word-type $F(1, 73) = 8.80$, $\eta_p^2 = .10$, $p < .01$ as participants recalled more stereotype-consistent descriptors (M = 46%) than stereotype-inconsistent descriptors (M = 40%). A main effect of WMC emerged $F(1, 73) = 6.45$, $\eta_p^2 = .08$, $p = .01$, wherein high, versus low, span persons recalled more descriptors (high WMC: 46%; low WMC: 40%). Finally, there was a reliable target-race x word-type interaction $F(1, 73) = 62.00$, $\eta_p^2 = .46$, $p < .01$, as persons recalled more stereotype-consistent, than –inconsistent, descriptors for the Black targets (stereotype-consistent:

[7] Ambiguous Targets could not be included in this analyses as the descriptor words were either Black- or Asian-(in)consistent and thus, could not be ambiguous consistent or inconsistent.
[8] Neutral words were not included in the analysis as there were half as many neutral, compared to stereotype-consistent or –inconsistent, words presented per list (2 versus 4), thus, the percentage of neutral words recalled is artificially inflated.

53%; stereotype-inconsistent: 30%) whereas the converse was true for Asian targets (stereotype-consistent: 38%; stereotype-inconsistent: 50%).

It is worth noting that a word-type x WMC trend emerged $F(1, 73) = 2.94$, $\eta_p^2 = .04$, $p = .09$. This finding was decomposed by running separate ANCOVAs for each rule set comparing error rate to only stereo-related lures for high versus low WMC persons. The results revealed that high WMC persons recalled a significantly greater number of stereotype-consistent descriptors (M = 51%) than low WMC persons (M = 41%), $F(1, 73) = 9.81$, $\eta_p^2 = .12$, $p < .01$, whereas all persons recalled a similar number of stereotype-inconsistent descriptors regardless of WMC, $F(1, 73) = .48$, p = ns, (high WMC: 41%; low WMC: 39%).

Turning to memory errors or intrusions, a main effect of word-type $F(1, 73) = 6.07$, $\eta_p^2 = .08$, $p = .02$ was found, as persons committed a greater number of stereotype-consistent intrusions ($M = .38$) than stereotype-inconsistent intrusions ($M = .19$). Additionally, the target-race x word-type interaction produced significant results such that persons committed a similar number of stereotype-consistent and –inconsistent intrusions for Black targets (stereotype-consistent: .22; stereotype-inconsistent: .25) but committed significantly more stereotype-consistent ($M = .53$) than –inconsistent intrusions ($M = .12$) for Asian targets.

Contrary to predictions, the main effect of WMC did not emerge. This null result was likely due to the low number of intrusions overall. Only 55 (of 77) participants committed any intrusions; 121 intrusions in total were recorded. When considering that there were a total of 462 lists recalled (77 participants, 6 lists each), this amounts to an average of .26 intrusion per list or 1.5 intrusions per person across all 6 lists.

Finally, my three a-priori predictions were tested. First, low WMC persons were expected to make more stereotype-consistent than –inconsistent errors. This prediction was not supported as low WMC persons committed a similar number of stereotype-consistent (M = .80) and – inconsistent (M = .46) intrusions, $F(1, 32) = 1.70$, p = ns. Next, low WMC persons were expected to commit more stereotype-consistent than neutral intrusions. Results supported this prediction as low WMC persons committed a greater number of stereotype-consistent (M = .80) than neutral (M = .03) intrusions, $F(1, 32) = 12.48$, $\eta_p^2 = .28$, $p < .01$. Finally, low, compared to high, WMC persons were expected to commit more stereotype-consistent intrusions. Persons committed an equal number of stereotype-consistent intrusions regardless of WMC (high WMC: .74; low WMC: .80), thus, the final prediction was not supported $F(1, 73) = .02$, p = ns.

Overall, the results partially supported predictions as high WMC persons had an advantage when it came to accurate recall. However, inconsistent with expectations low WMC persons recalled fewer stereotype-consistent descriptors than high WMC; whereas they were expected to recall fewer stereotype-inconsistent descriptors than their high WMC counterparts. Potentially, high WMC persons were able to use the targets implied ethnicity to act as a category label which aided in stereotype-consistent recall. Support for this suggestion comes from the fact that high and low WMC persons recalled a similar number of descriptors when the target was ambiguous (high WMC: 42%; low WMC: 37%), $F(1, 73) = 3.09$, p = ns. Importantly, this same group of people did not commit a greater number of stereotype-consistent intrusions, or any type intrusions for that matter. However, these findings failed to support the suggestion that low WMC persons are more susceptible to making stereotypical intrusions on retrieval tasks than 1) stereotype-inconsistent intrusions, or 2) high WMC persons.

4.3.4 Cued Recall Task (A2) First, considering cued recall accuracy (attributing the correct target to a listed descriptor), a Target Race (Black, Asian) x 2 WMC (high, low) mixed-model ANCOVA was run with WMC as a between-subjects variable. Black American and Asian

American prejudice levels were entered as control variables. There were no significant effects or interactions. Next, cued recall error was evaluated (attributing the incorrect target to a listed descriptor). A mixed-model ANCOVA was run: 2 Error-Type x 2 Target Race x 2 WMC. Error-type (stereotype-consistent, -inconsistent) and Target Race (Black, Asian) served as the within-subjects factor whereas WMC (high, low) was between-subjects. Only a main effect of error-type emerged as all persons committed a greater number of stereotype-consistent errors (M = 20%) than –inconsistent errors (M = 17%) regardless of WMC, $F(1, 73) = 8.46$, $\eta_p^2 = .10$, $p < .01$ (See Tables 7 and 8 for overall accuracy and error rates). These results suggest that WMC has no effect on performance for cued recall (source monitoring) tasks.

4.3.5 Familiarity Task (B) Low and high WMC persons were expected to perform equally well netting a similar number of hits and false alarms. To test this prediction a mixed-model ANCOVA was run on the overall design: 2 Target-Race x 3 Word-Type x 2 WMC. Target-race (Black, Asian)[9] and word-type (stereotype-consistent, -inconsistent, neutral) served as the within-subject factors while WMC (high, low) was between-subjects. Dependent variables of interest include hits and false alarms. Black American and Asian American prejudice levels were entered as control variables. See Tables 9 and 10 for overall recall/errors for all persons and all word types. No significant main effects or interactions emerged for hits. Only a main effect of target-race was significant for false alarms $F(1,62) = 6.01$, $\eta_p^2 = .09$, $p = .02$ as persons made more false alarms to Black targets (5%) than Asian targets (2%). These results support the predictions that: 1) low and high WMC persons would perform equally well netting similar hit and false alarm rates, and 2) all persons would recognize all words (stereotype-consistent, -inconsistent, and –neutral) with equal accuracy.

[9] Ambiguous Targets could not be included in this analyses as the descriptor words were either Black- or Asian-(in)consistent and thus, could not be ambiguous consistent or inconsistent.

4.4 Discussion

The current findings provide some support for the suggestion that WMC affects the influence of stereotypical information when performing retrieval tasks. The effect of WMC during list intrusions was limited but this result was likely due to the extremely low intrusion rate. A low intrusion rate for list recall is consistent with previous research (Kane & Engle, 2000; which is not unusual for interference studies, see Melton & Irwin, 1940). However, WMC did affect accurate list retrieval; but only stereotype-consistent retrieval such that high, versus low, WMC persons recalled more of this type of descriptor whereas all persons recalled a similar number of stereotype-inconsistent descriptors. Although this finding is different than the predicted finding wherein low, compared to high, WMC persons were expected to recall fewer stereotype-inconsistent descriptors; potentially, high WMC persons had the available capacity to adopt the ethnic target names as akin to a category label, thus, providing a schema that aided in schema-consistent recall.

The present findings do lend some support and extend the retrieval portion of Unsworth and Engle's (2007) maintenance and retrieval theory of individual differences in WMC as WMC was found to influence performance on retrieval (i.e., accurate recall) but not familiarity tasks. Previous researchers suggest that the differential in accurate list recall between high and low WMC persons is due to one's susceptibility to proactive interference (Kane and Engle, 2000). Proactive interference (PI) occurs when retrieval of a current list (i.e., List 3) is impaired due to the prior study of a similar list (i.e., List 2). Consistent with memory theories suggesting that the avoidance of PI requires controlled processing (e.g., Anderson & Bjork, 1994), low WMC persons were found to recall fewer and fewer accurate items as the list presentations progressed (Kane and Engle, 2000). In an effort to replicate this finding, accurate recall was calculated separately for each of the four lists with ethnic targets.[10] I was unable to demonstrate the effects of PI as accurate recall did not decrease across lists generally (see Figure 4) or across lists for any single word type (see Figures 5-7). Potentially, the effect did not emerge as the lists in the current investigation were necessarily different than the lists used by Kane and Engle (2000) as the current lists 1) contained a target name that could be utilized as a category label and 2) included subcategories of words (stereotype-consistent, -inconsistent, and –neutral). Conversely, Kane and Engle presented participants with multiple lists of animals, no category labels, and no word subcategories. As differences in accurate recall emerged as a function of individual WMC regardless of the lack of evidence supporting PI, the significance of the present finding is twofold. First, I suggest that this result lends further support to supposition that it was in fact the spontaneous encoding of the targets ethnicity that allowed for a superior performance of the high versus the low WMC persons. Second, the current results may better be explained by other models of individual differences in working memory capacity (see the General Discussion more further explanation).

[10] As accurate list recall was less for the ethnically ambiguous targets these lists were not included in the current comparisons.

CHAPTER 5:

GENERAL DISCUSSION

This study represents one of the first attempts to establish connections between individual WMC and stereotype use. The overall aim of the present investigation was to test whether individual WMC impacted reliance on stereotypic information during maintenance and retrieval tasks. Stereotypical associations are automatic and cognitively efficient, thus, without sufficient resources to suppress, low relative to high WMC persons were expected to make more stereotype-consistent errors. Conversely, all participants were expected to perform similarly on inhibition and familiarity tasks as WMC has not been a factor in these types of tasks (Unsworth & Engle, 2007). Results partially supported these predictions.

The results from the pilot study support the suggestion that all persons regardless of individual WMC maintain similar knowledge of common social stereotypes related to African American and Asian American men. Specifically, 50% or more of persons agreed that 91% of the original person descriptors were common social stereotypes of African American (e.g., athletic, loud) and Asian American men (e.g., intelligent, short). This result allowed me to move forward in testing my primary hypotheses with the confidence that high and low WMC persons maintain similar knowledge of common racial stereotypes.

The goal of Experiment 1 was to assess the effect of WMC on stereotype use during cognitive maintenance versus inhibition tasks. The results revealed that low WMC persons were more susceptible to making stereotypical errors on maintenance tasks than 1) other errors (i.e., neutral or stereotype-inconsistent), or 2) high WMC persons; no performance differences were observed between high and low capacity persons on inhibition tasks. It is important to note that individual prejudice was controlled for during all analyses. This result is significant as it provides support for the maintenance portion of Unsworth and Engle's (2007) maintenance and retrieval theory of individual differences in WMC and extends extant literature to suggest that in fact overreliance on stereotypic information may be affected by individual WMC putting low WMC persons at greater risk of committing stereotype-consistent errors than their high WMC counterparts. Clearly the suggestion is not that low WMC persons are likely to commit stereotype-consistent errors during any decision-making task as performance was equivalent for all persons during the inhibition task.

The modified go/no-go paradigm used in the present investigation is not the only example of a maintenance task. A maintenance task is any task that requires active maintenance of task relevant information in order to override fairly automatic responses. Based on the current results, I suggest that low, compared to high WMC persons are more susceptible to making stereotypical errors on any maintenance task. Furthermore, as ethnic stereotypes are only one example of a stereotype, low WMC persons are likely vulnerable to the undue influence of other stereotypes when performing maintenance tasks.

The goal of Experiment 2 was to assess the effect of WMC on stereotype use during cognitive retrieval versus familiarity tasks. Overall, high WMC persons recalled a greater number of person descriptors than low WMC persons. More specifically, they recalled a greater number of stereotype-consistent person descriptors than low WMC persons; all persons recalled a similar number of stereotype-inconsistent descriptors. I suggest that this finding was the result of high WMC persons utilizing their additional capacity to process the target-name as a category label which provided them with the advantage when recalling category-consistent words. It is important to note that this advantage during accurate recall did not lead to a disadvantage during inaccurate recall (i.e., intrusions) as all persons committed a similar number of intrusions.

Moreover, this suggests that high WMC persons' capacity was not compromised by their ability to label and store the stereotyped information.

I originally expected that stereotypes would influence accurate descriptor recall as a function of WMC such that low WMC persons would recall fewer stereotype-inconsistent than –consistent words while descriptor stereotypicality was not expected to affect high WMC accurate word recall. Instead, high WMC persons recalled more stereotype-consistent than –inconsistent descriptors while low WMC persons recalled a similar number of descriptors regardless of stereotype-consistency. Although different than expected, this result still supports two important ideas. First, high, versus low, WMC provides an advantage with regard to accurate list recall. Second, stereotypes do in fact differentially affect high and low WMC persons' list recall.

The list presentation/recall paradigm used in the present investigation is not the only example of a retrieval task. A retrieval task is any task that requires retrieving task-relevant information in the presence of task-irrelevant information. Based on the current results, I suggest that stereotypical information may benefit high, compared to low, WMC persons on any retrieval task. Moreover, any task that makes use of a template or schema of sorts will benefit the high WMC person, not just stereotypes but any semantic category where they can code and store the information more efficiently than a low WMC person

5.1 Implications

5.1.1 Individual Differences in Stereotype Use

A myriad of factors have been put forth to explain individual differences and the influence of stereotypes on behavior. For example, Christiansen, Kaplan, and Jones (1999) found that individual differences in prejudice affected participants' ratings of an ostensible application for a minority scholarship such that more prejudice persons rated the scholarship containing a single negative stereotype more negatively than less prejudice persons. Glaser and Knowles (2008) found that persons rating high in implicit negative attitude toward prejudice demonstrated no correlation between the race/weapon stereotype (associating Black men and guns) and the shooter bias (tendency to shoot armed Black men faster than armed White men). Additionally, an implicit belief that oneself was prejudiced alone was not correlated with the race/weapon stereotype-shooter bias relationship. However, persons believing themselves to be prejudiced and rating low in implicit negative attitude toward prejudice (i.e., not thinking prejudice was bad) achieved the greatest correlation between the race/weapon stereotype and shooter bias. Research by Gushue and Carter (2000) revealed that heighted anxiety about race was linked to better memory for stereotype-inconsistent information. As touched on here, most of the research looking at the effect of individual differences on stereotype-use focuses on the influence of various forms of prejudice. It is interesting to note that in the present investigation individual WMC affected stereotype-use regardless of individual prejudice level. That is, even after controlling for individual prejudice level in all Experiment 2 and 3 analyses, WMC still impacted maintenance and retrieval task performance. This result suggests that WMC may be more impactful in determining when stereotypes influence behavior than prejudice. That is, with sufficient capacity high WMC persons were able to avoid making stereotype-consistent errors during a maintenance task regardless of their individual prejudice level.

5.1.2 Executive Control and Stereotype Use

Given that a) stereotypes are ingrained associations that are automatically activated and b) working memory is needed in order to override automatic response biases; it is surprising that little research has been conducted examining the relationship between working memory and stereotype use (see Kleider, et al., 2012 for an exception). One such study conducted by Payne (2005), found a link between engaged executive control and stereotype use. Although relevant to the current work, what this study does not provide is a link between one's static available capacity and the influence of automatically activated stereotypes. One's engaged control is not synonymous with one's available capacity. A low WMC person has the capacity to engage executive control; it is under certain circumstances (maintenance and retrieval tasks) that their capacity is overwhelmed and they succumb to automatic processing. Thus, the performance differences observed by Payne (2005) may not be related to individual differences in WMC at all. A second difference noted between the findings in the present study and those of the Payne study is task-type. Payne observed performance differences in weapon identification, word evaluation and person-judgment tasks as a function of engaged executive control. None of these tasks could be classified as maintenance or retrieval tasks. Drawing on Unsworth and Engle's (2007) maintenance and retrieval theory of individual differences in WMC, I would suggest that the differences found in Payne (2005) due to engaged executive control are distinct and not necessarily related to individual available capacity (i.e., WMC).

5.1.3 Theory

The current work supports and extends Unsworth and Engle's (2007) maintenance and retrieval theory of individual differences in WMC by demonstrating a link between individual WMC and stereotype use. However, there were some distinctions between predictions made by the theory and the implications of the present results. Specifically, high WMC persons recalled more list items in both the current study and the Kane and Engle (2000) study but the underlying cognitive mechanism was different. In the current task, there was no evidence of proactive interference; instead, I suggest that the nature of the list (describing an ostensible person) encouraged each list to be viewed as distinct preventing the buildup of proactive interference. Potentially, each set of descriptors could be chunked to form a mental image of the described, this image was likely quite distinct from ‘person’ to ‘person,’ thus, descriptors used in list 2 did not seem relevant to list 3.

The lack of evidence for proactive interference does not explain the advantage that high WMC persons demonstrated over low WMC persons when recalling stereotype-consistent words. I suggest that this advantage was the result of the semantic or semantically related list information which allowed high WMC people to activate a heuristic or template leading to efficient storage. Low WMC people do not have the required capacity to spontaneously categorize and store this information as efficiently. Thus, high WMC persons recalled significantly more stereotype-consistent descriptors than low WMC persons. Conversely, the list items in the Kane and Engle (2000) study included animals which are categories but not sub-categories that tie to a heuristic or template or rule; and thus, did not present such an opportunity for this sorting or subgrouping.

An alternative explanation for the retrieval results comes from the binding theory of individual difference in WMC (Oberauer, Süß, Wilhelm, & Sander, 2007) wherein the assigned function of WM is to establish and maintain ad hoc bindings between representations. Under this framework high WMC persons are better able to temporarily bind stimuli with response than low WMC persons. However, one would expect, if increased list recall was just a matter of superior binding, then the stereotype-consistency of the words should not affect high WMC persons retrieval performance. That is, high WMC persons should have had the advantage with both stereotype-consistent and –inconsistent words if their increased list retrieval was simply a function of stronger stimulus-response binding than low WMC persons. Potentially the suggested labeling/storage benefit afforded to high WMC persons is underpinned by categorization or semantic storage.

There is much debate regarding how categorical knowledge is represented. Specifically, prototype theories suggest that category membership is deteremined based on comparing similarities of a new item to the average of one‘s exemplar experience whereas exemplar theory argues that membership is instead deteremined by comparing to specific representations encountered during personal training (Posner & Keele, 1968, 1970; Medin & Schaffer, 1978; Smith & Minda, 2000). Applying this to the findings in the current study, potentially, high WMC persons with their superior processing abilities, encoded target names extrapolating implied ethnicity, which activated general knowledge regarding a representative ethnic member. A prototype, as developed by cognitive psychologist Eleanor Rosch (1975), assumes the existance of perceptual salience impacting degrees of category membership such that there are central and marginal members. Prototypes are considered to be psychologically real (see Rosch, 1978). A stereotype (Putnam, 1975) is seen as frame-and-script based information which provides the context for a prototype representation. Simply put, the notion of stereotype is a sociocentric one (see Silverstein, 2004, 2007). Regardless of the specific distinctions, it is important to note that the two constructs are closely connected. Thus, prototype activation facilitates stereotype activation and similarly stereotype activation facilitates prototype activation. Conversely, the link between exemplars and stereotypes is much more tenuous. That is, activating an exemplar category member who maintains various qualities, some stereotype-consistent some – inconsistent, should promote equal retrieval of all descriptors. Thus, the present results may best be explained by a prototype theory of categorization. Moreover, these findings suggest that prototype versus exemplar categorization was the automatic default.

In summation, the present investigation extends extant literature regarding individual differences that contribute to stereotype use. It is among the first to link individual working memory capacity to stereotype use (See Kleider, Knuycky, & Cavrak, 2012). And, it lends some support to while extending Unsworth and Engle's (2007) maintenance and retrieval theory of individual differences in WMC.

5.1.4 WMC, Stereotypes, and Memory Bias

The present retrieval result wherein high WMC persons accurately recalled more stereotype-consistent list descriptors than low WMC persons may lead one to suspect that high WMC persons are at an increased risk of biased decision making. However, the findings from Kleider and colleagues (2012) suggest that this is not the case. As previously described, Kleider and colleagues had high and low WMC persons act as mock-jurors and render verdict decisions for Black and White defendants. It is reasonable to surmise that the participants in this study attempted to retrieve case facts for review prior to rendering a verdict. Thus, the current results would suggest that high WMC persons would likely retrieve more stereotype-consistent case facts (e.g., the Black defendant was aggressive) which in turn, would bias their decisions. However, actual results revealed that high WMC persons rendered similar verdict decisions regardless of defendant race while low WMC persons assigned more guilty verdicts to Black than White defendants in the cognitive load condition. From this, I suggest that high WMC persons maintain more information generally, and more stereotype-consistent information specifically, than low WMC persons; however, their ability to retain that information does not bias their decisions. Instead high, compared to low, WMC persons are able to apply their superior capacity to making more consistent judgments.

5.3 Future Directions

Participants in the retrieval task generated surprisingly few list intrusions. Although this result is common for this type of task, Kane and Engle (2000) suggested that fluency tasks provide a rich opportunity for intrusions. Potentially, such a task could be used to test whether individual WMC affects stereotypical intrusion rate. The expectation would be that low WMC persons would generate significantly more list intrusions than high WMC persons. Testing this prediction would be a worthwhile endeavor as the implications of such a result would be significant. A tendency to recall any misinformation is of interest, however, a tendency to disproportionately recall stereotype-consistent misinformation likely informs one's propensity to make biased decisions.

As previously mentioned, finding that high WMC persons recalled more stereotype-consistent descriptors than either 1) stereotype-inconsistent descriptors or 2) low WMC persons was unexpected. However, potentially was the result of high WMC persons spontaneously encoding the target-name which they were able to use as akin to a category label facilitating later category-consistent recall. This suggestion is easily tested by inquiring about persons' memory/impression of the target after list recall. I would expect high, versus low, WMC persons to have better memory of the ethnic target names and/or automatically encode an inferred ethnicity. Additionally, the present results suggest that it was high WMC persons' superior capacity that allowed them to encode and extrapolate target information. However, if this information was more readily available (i.e., automatically encoded), then it should promote the automatic stereotype-consistent associations for low WMC persons too, allowing for similar retrieval performance across groups. Specifically, this suggestion could be tested by presenting an ethnically prototypical category member with each descriptor list. Under these conditions, I would expect high and low WMC persons to accurately retrieve a similar number of descriptors and for all persons to retrieve more stereotype-consistent than –inconsistent descriptors.

As discussed above, the present investigation generated support for the prototype versus exemplar theory of categorization. Future studies could investigate whether metacognition could modify categorization techniques. Specifically, if category membership was made explicit via the presentation of target demographics or photographs, in lieu of names and participants were instructed to activate category exemplars in an effort to aid in retrieval, it would be interesting to note whether persons' stereotype-inconsistent retrieval increased. An increase in stereotype-inconsistent descriptor retrieval under these conditions would support a potential method for decreasing bias. Potentially, WMC would differentially affect stereotype-descriptor retrieval under these conditions, such that high WMC persons retrieve more descriptors overall with a similar rate of stereotype-consistent and –inconsistent recall (supporting exemplar categorization) whereas low WMC persons recall more stereotype-consistent than –inconsistent descriptors (as encoding the target race has been made automatic). This result would suggest that exemplar- versus prototype-based categorization requires increased capacity.

Finally, efforts should be made to extend the present findings as stereotypes are but one example of a heuristic or mental shortcut that creates an automatic response bias. Other examples of heuristics include the fluency heuristic which suggests that items made faster to process through priming (semantic e.g., Whittlesea, 1993, or perceptual e.g., Jacoby & Whitehouse, 1989), are more likely to be recognized as 'old.' The distinctiveness heuristic, when tested with recognition memory often involves presenting participants with a list of words or a list of words coupled with a picture for later recall. All participants recall true 'old' words with similar accuracy but participants shown the word/picture list use the distinctive presentation to avoid false alarms to lure items (e.g., Dodson & Hege, 2005). According to Tversky & Kahneman's (1974) availability heuristic, people make decisions based on the ease with which something comes to mind. For example, one may guess how common heart attacks are among middle aged Americans by determining the percentage of their acquaintances that fit the description. Inferring form the present investigation, it is reasonable to suggest that individual WMC may differentially impact one's reliance on such heuristics when under pressure. Specifically, low WMC persons may be at increased risk of committing heuristic-consistent errors whereas high WMC persons may be able to spontaneously adopt the heuristic as an aid to facilitate task performance.

5.4 Conclusion

The present research suggests that sufficient individual working memory capacity may be a prerequisite to 1) avoiding the influence of automatically activated stereotypes on behavior when performing maintenance-type tasks and, 2) spontaneously incorporating heuristic-consistent information into one's cognitive toolbox to aid in performing retrieval-type tasks. This suggestion is significant as stereotypes have been found to influence many behaviors including categorization (Blair & Banaji, 1996, Experiment 2), memory (Stewart, Weeks, & Lupfer, 2003), and decision making (Payne, Lambert, and Jacoby, 2002), however, to date, limited research has identified WMC as a factor in this relationship (only Kleider, et al., 2012).

REFERENCES

Allport, G. W. (1954). *The nature of prejudice.* Reading, MA: Addison-Wesley.

Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in long-term memory: A new taxonomy. In D. Dagenbach and T Carr (Eds.) *Inhibitory processes in attention, memory, and language* (pp. 265-325). San Diego, CA: Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation.* New York: Academic Press.

Baddeley, A. D. (1992). Working memory. *Science, 255,* 556-559.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A,* 5-28.

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

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (pp. 47-89). New York: Academic Press.

Baddeley, A. D., Lewis, V. F. J., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology, 36,* 233-252.

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology, 70,* 1142-1163.

Bogardus, E. S. (1925). Measuring social distance. *Journal of Applied Sociology, 9,* 299-309.

Brandimonte, M. A., Hitch, G. J., & Bishop, D. V. M. (1992). Influence of short-term memory codes on visual image processing: Evidence from image transformation tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 157-165.

Brewer, M. B. (1988). A dual process model of impression formation. In T. Srull and R. Wyer (Eds.), *Advances in social cognition* Vol. 1, Earlbaum.

Bunting, M. F., Conway, A. R. A., & Heitz, R. P. (2004). Individual differences in the fan effect and working memory capacity. *Journal of Memory and Language, 51,* 604-622.

Christiansen, N. D., Kaplan, M. F., & Jones, C. (1999). Racism and the social judgment process: Individual differences in the use of stereotypes. *Social Behavior and Personality, 27,* 129-144.

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

Cramer, P., & Steinwert, T. (1998). Thin is good, fat is bad: How early does it begin? *Journal of Applied Developmental Psychology, 19,* 429-454.

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

Delaney, P. F., & Sahakyan, L. (2007). Unexpected costs of high working memory capacity following directed forgetting and contextual change manipulations. *Memory and Cognition, 35,* 1074-1082.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5-18.

Dodson, C. S. & Hege, A. C. G. (2005). Speeded retrieval abolishes the false-memory suppression effect: Evidence for the distinctiveness heuristic. *Psychonomic Bulletin and Review, 12,* 726-731.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348-368.

Fiske, S. T. (2004). Social beings: A core motives approach to social psychology. Hoboken, NJ: Wiley.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1-74). New York Academic Press.

Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice. *Journal of Experimental Social Psychology, 44,* 164-172.

Goldinger, S. D., Kleider, H. M., Azuma, T., & Beike, D. R. (2003). "Blaming the victim" under memory load. *Psychological Science, 14,* 81-85.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment, *Neuron, 44,* 389-400.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74,* 1464-1480.

Gushue, G. V., & Carter, R. T. (2000). Remembering race: White racial identity attitudes and two aspects of social memory. *Journal of Counseling Psychology, 47,* 199-210.

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General, 118,* 126-135.

Jones, C. S., & Kaplan, M. F. (2003). The effects of racially stereotypical crimes on juror decision-making and information-processing strategies. *Basic and Applied Social Psychology, 25,* 1-13.

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

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

Kahneman, D. (1995). Varieties of counterfactual thinking. In N. J. Neal and J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking.* Hilldale NJ.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58,* 697-720.

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,* 169-183.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R.W. (2007). Variation in working memory capacity as variation in executive attenriona and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse(Eds.), *Variation in working memory.* New York: Oxford University Press.

Kane, M. J., & Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long-term retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 333-358.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47-70.

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

Kleider, H. M., Knuycky, L. R., & Cavrak, S. E. (2012). Deciding the fate of others: The cognitive underpinnings of racially biased juror decision making. *Journal of General Psychology, 139,* 175-193.

Kleider, H.M., Pezdek, K., Goldinger, S. D., & Kirk, A. (2008). Schema-driven source misattribution errors: Remembering the expected from a witnessed event. *Applied Cognitive Psychology, 22,* 1-20.

Klein, K. & Boals, A. (2001). The relationship of life event stress and working memory capacity. *Applied Cognitive Psychology, 15,* 565-579.

Kurtzman, H. (1985). *Studies in syntactic ambiguity resolution*. Doctoral dissertation, MIT, Cambridge, MA. Distributed by Indiana University Linguistics Club, Bloomington.

Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General, 130,* 199-207.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology, 24,* 56-98.

Macmillian, N. A. & Creelman, C. D. (2005). Detection Theory: A User's Guide (2nd Ed.). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology, 66,* 37-47.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85,* 207-238.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63,* 8-97.

Moore, A. B., Clark, B. A., & Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science, 19,* 549-557.

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

Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General, 134,* 368-387.

Oberauer, K, Süß, H-M, Wilhelm, O., & Sander, N. (2007). Individual differences in working memory capacity and reasoning ability. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse(Eds.), *Variation in working memory.* New York: Oxford University Press.

Oliver, M. B. (1999). Caucasian viewer's memory of Black and White criminal suspects in the news. *Journal of Communication, 49,* 46-60.

Oliver, M. B., & Fonash, D. (2002). Race and crime in the news: Whites' identification and misidentification of violent and nonviolent criminal suspects. *Media Psychology, 4,* 137-156.

Payne, B. K. (2005). Conceptualizing control in social cognition: How executive function modulates the expression of automatic stereotyping. *Journal of Personality and Social Psychology, 89,* 488-503.

Payne, B. K., Lambert, A. J., & Jacoby, L. L. (2002). Beat laid plans: Effects of goal on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology, 38,* 384-396.

Payne, B. K., Shimizu, Y., & Jacoby, L. L. (2005). Mental control and visual illusions: Toward explaining race-based weapon misidentifications. *Journal of Experimental Social Psychology, 41,* 36-47.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353-363.

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology, 83,* 304-308.

Putnam, H. (1975). The meaning of meaning, in his *Philosophical Papers, Volume 2: Mind, Language, and Reality* (Cambridge: Cambridge University Press), pp 215-257.

Redick, T. S., Calvo, A., Gay, C. F., & Engle, R. W. (2011). Working memory capacity and go/no-go task performance: Selective effects of updating, maintenance, and inhibition. Journal of Experimental Psychology: Learning, Memory, and Cognition, 37, 308-324.

Richeson, J.A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science, 14,* 287-290.

Ricker, J. H., & Axelrod, B. N. (1994). Analysis of an oral paradigm for the Trail Making Test. *Assessment, 1,* 47-51.

Rosch, E. (1975). The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 303-322.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), Cognition and Categorization (hillside, NJ: Earlbaum), pp. 28-48.

Rosen, V. M., Bergeson, J. L., Putnam, K., Harwell, A., & Sunderland, T. (2002). Working memory and apolipoprotein E: What's the connection? *Neuropsychologia, 40,* 2226-2233.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85,* 440-452.

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General, 136,* 414-429.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology, 22,* 261-273.

Silverstein, M. (2004). 'Cultural' concepts and the language-culture nexus. Current Anthropology, 45, 621-652.

Silverstein, M. (2007). How knowledge begets communication begets knowledge: Textuality and contextuality in knowing and learning. Intercultural Communication Review, 5, 36-60.

Smith, J. D. & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 3-27.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74,* 1-29.

Stewart, T. L., Weeks, M., & Lupfer, M. B. (2003). Spontaneous stereotyping: A matter of prejudice? *Social Cognition, 21,* 263-298.

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

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and Biases, *Science, 185,* 1124-1131.

Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language, 54,* 68-80.

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

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., Redick, T. S., Heitz, R. P., Broadway, J. M., & 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.

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1235-1253.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441-517.

APPENDIX

Table 1. List of person-descriptors presented to participants (Pilot Experiment) in an effort to identify common racial stereotypes. Marks ('x' or '0') under a category indicate whether a descriptor was originally predicted to be Black stereotype-consistent, Asian stereotype-consistent, or neither. The 'x' indicates that the prediction was supported while the '0' indicates that is was not.

	Black	Asian	Neutral
Aggressive	X		
Amusing	X		
Approachable			0
Assertive	X		
Athletic	X		
Attractive			0
Basketball Player	X		
Brainy		X	
Brave			0
Bright		X	
Calm		X	
Caring			X
Comical	X		
Compassionate			0
Con	X		
Conscientious		0	
Conventional		0	
Convict	X		
Courageous			0
Criminal	X		
Crook	X		
Culprit	X		
Customary		X	
Daring			0
Delinquent	X		
Destructive	X		
Diligent		X	
Dim	0		
Disciplined		X	
Easygoing		0	
Enduring		0	
Felon	X		
Fetching			0
Flashy	X		
Football Player	X		
Forceful	X		
Friendly			0
Funny	X		
Good at Math		X	
Good Looking			0
Gorgeous			0
Gutsy			0
Hardworking		X	

Helpful			**0**
Hilarious	**X**		
Hostile	**X**		
Humble		**X**	
Humorous	**X**		
Industrious		**X**	
Intelligent		**X**	
Kind			**0**
Lazy	**X**		
Linebacker	**X**		
Little		**X**	
Looker			**0**
Loud	**X**		
Meek		**X**	
Meticulous		**X**	
Mild		**X**	
Modest		**X**	
Noisy	**X**		
Offender	**X**		
Old-Fashioned		**X**	
Outgoing			**0**
Patient		**X**	
Persistent		**X**	
Petit		**X**	
Point guard	**X**		
Poor Student	**X**		
Productive		**X**	
Promiscuous	**X**		
Quarterback	**X**		
Quiet		**X**	
Relaxed		**X**	
Reserved		**X**	
Rowdy	**X**		
Running back	**X**		
Self-Controlled		**X**	
Sexual	**X**		
Sharp		**X**	
Short		**X**	
Shy		**X**	
Small in Stature		**X**	
Smart		**X**	
Sociable			**0**
Soft spoken		**X**	
Stupid	**X**		
Submissive		**X**	
Thoughtful			**0**
Tolerant		**0**	
Traditional		**X**	
Understanding			**X**
Unintelligent	**X**		
Unlawful	**X**		
Unmotivated	**X**		
Unproductive	**X**		
Violent	**X**		
Welcoming			**0**
Witty	**0**		

Table 2. Complete list of stereotype-consistent descriptors and participant agreement ratings.

African American Consistent Descriptors

Descriptor	African American	Asian
Aggressive	91%	3%
Amusing	57%	8%
Assertive	60%	22%
Athletic	97%	0%
Basketball Player	98%	1%
Comical	69%	8%
Con	75%	1%
Convict	88%	0%
Criminal	88%	0%
Crook	84%	3%
Culprit	65%	5%
Delinquent	73%	13%
Destructive	71%	10%
Felon	88%	10%
Flashy	82%	2%
Football Player	98%	0%
Forceful	77%	7%
Funny	69%	3%
Hilarious	70%	4%
Hostile	70%	8%
Humorous	71%	5%
Lazy	83%	2%
Linebacker	95%	1%
Loud	91%	0%
Noisy	83%	2%
Offender	77%	1%
Point guard	82%	1%
Poor Student	78%	2%
Promiscuous	75%	3%
Quarterback	92%	0%
Rowdy	88%	1%
Running back	92%	1%
Sexual	83%	0%
Stupid	68%	2%
Unintelligent	67%	2%
Unlawful	78%	2%
Unmotivated	70%	3%
Unproductive	71%	4%
Violent	85%	2%

Asian Consistent Descriptors

Descriptor	African American	Asian
Brainy	3%	94%
Bright	9%	79%
Calm	3%	71%
Customary	3%	63%
Diligent	8%	72%
Disciplined	10%	83%
Good at Math	3%	96%
Hardworking	24%	65%
Humble	14%	57%
Industrious	18%	56%
Intelligent	17%	79%
Little	2%	92%
Meek	6%	52%
Meticulous	11%	52%
Modest	8%	65%
Old-Fashioned	13%	61%
Patient	5%	57%
Petit	2%	84%
Productive	14%	76%
Quiet	3%	87%
Reserved	2%	87%
Sharp	24%	56%
Self-Controlled	8%	67%
Short	2%	92%
Shy	3%	79%
Small in Stature	3%	93%
Smart	5%	92%
Soft Spoken	2%	80%
Submissive	12%	50%
Traditional	7%	81%
Uncomplaining	3%	57%

Table 3. Final set of descriptors used for memory test.

African American Consistent Descriptors

Descriptor	African American	Asian
Aggressive	91%	3%
Amusing	57%	8%
Assertive	60%	22%
Athletic	97%	0%
Basketball Player	98%	1%
Comical	69%	8%
Convict	88%	0%
Criminal	88%	0%
Delinquent	73%	13%
Destructive	71%	10%
Flashy	82%	2%
Football Player	98%	0%
Forceful	77%	7%
Funny	69%	3%
Hilarious	70%	4%
Humorous	71%	5%
Lazy	83%	2%
Linebacker	95%	1%
Loud	91%	0%
Noisy	83%	2%
Promiscuous	75%	3%
Quarterback	92%	0%
Rowdy	88%	1%
Running back	92%	1%
Sexual	83%	0%
Unintelligent	67%	2%
Unlawful	78%	2%
Unmotivated	70%	3%
Unproductive	71%	4%
Violent	85%	2%

Asian Consistent Descriptors

Descriptor	African American	Asian
Brainy	3%	94%
Bright	9%	79%
Calm	3%	71%
Customary	3%	63%
Diligent	8%	72%
Disciplined	10%	83%
Good at Math	3%	96%
Hardworking	24%	65%
Humble	14%	57%
Industrious	18%	56%
Intelligent	17%	79%
Little	2%	92%
Meticulous	11%	52%
Mild	8%	49%
Modest	8%	65%
Old-Fashioned	13%	61%
Patient	5%	57%
Petit	2%	84%
Productive	14%	76%
Quiet	3%	87%
Reserved	2%	87%
Self-Controlled	8%	67%
Short	2%	92%
Shy	3%	79%
Small in Stature	3%	93%
Smart	5%	92%
Soft Spoken	2%	80%
Submissive	12%	50%
Traditional	7%	81%
Uncomplaining	3%	57%

Table 4. Complete list of stereotype-neutral descriptors. Marks ('x' or '0') under a category indicate whether a descriptor was used in the memory test. The 'x' indicates that the descriptor was used while the '0' indicates that is was not.

	Neutral
Bowls	X
Chews Gum	X
Dislikes Cereal	X
Dislikes Junk Mail	X
Drinks Coffee	X
Drinks Water	X
Drives Fast	0
Eats Bananas	X
Eats Meat	0
Enjoys Rain	X
From the City	0
Gardens	0
Grills Burgers	0
Has Close Friends	X
Has Pets	X
Has Short Nails	0
Has Vivid Dreams	X
Has a Brother	X
Hates Traffic	0
Is 28	X
Likes Blue	X
Likes Good Food	0
Likes Kids	0
Likes to Draw	0
Partial to Green	X
Plays Angry Birds	X
Rides a Bike	X
Surfs the Internet	0
Takes Hot Showers	0
Wants to Retire	0
Watches Movies	X
Watches TV	0
Watches the News	0
Wears Cologne	0
Wears Jeans	0

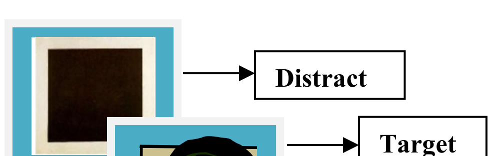

Figure 1. Illustration of the trial types in the conditional go/no-go task (Experiment 1A). For the present example, the rule set was stereotype-consistent and required a go-response to Black faces, weapons, and circles.

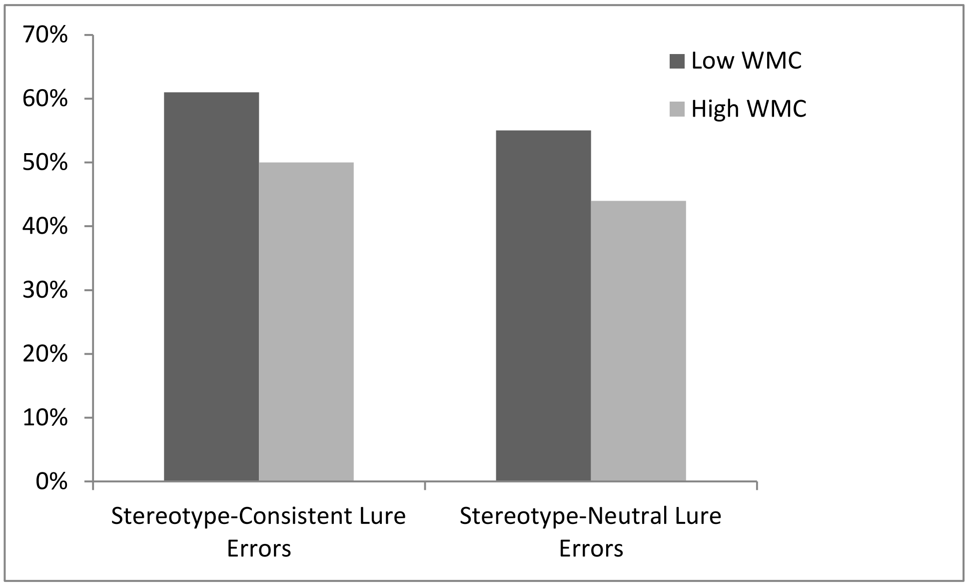

Figure 2. Maintenance Task Performance: Errors made to stereotype-consistent versus stereotype-neutral lures as a function of WMC.

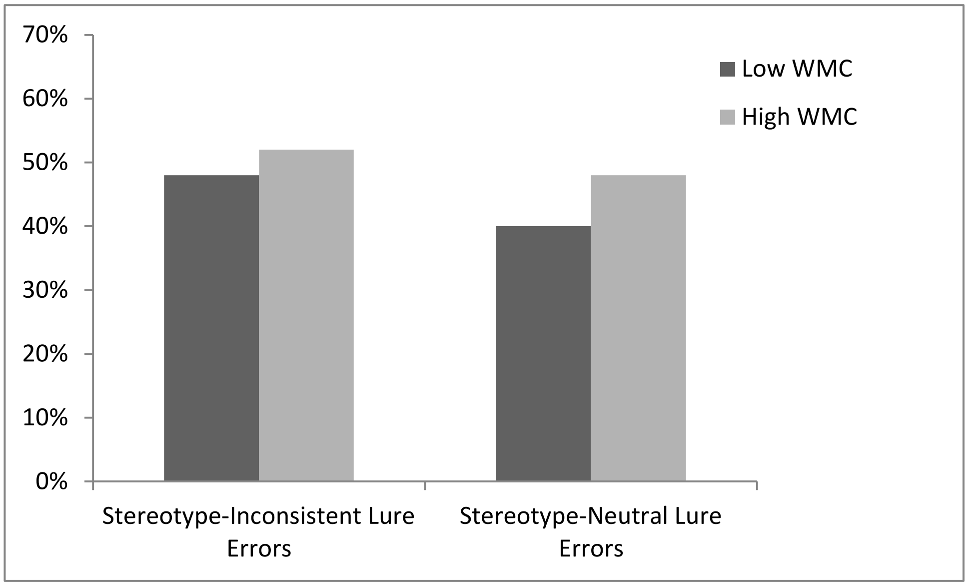

Figure 3. Maintenance Task Performance: Errors made to stereotype-inconsistent versus stereotype-neutral lures as a function of WMC.

Table 5. Retrieval Task Performance: Accurate recall as a function of target race, word type, and WMC.

Target Race	Word Type	WMC	Accurate Recall %
		Low WMC	49%
	Stereotype-Consistent	High WMC	58%
Black		Low WMC	29%
	Stereotype-Inconsistent	High WMC	31%
		Low WMC	33%
	Stereotype-Consistent	High WMC	43%
Asian		Low WMC	50%
	Stereotype-Inconsistent	High WMC	51%

.

Table 6. Retrieval Task Performance: Inaccurate recall (i.e., intrusions) as a function of target race, word type, and WMC.

Target Race	Word Type	WMC	Intrusion Count
		Low WMC	0.29
	Stereotype-Consistent	High WMC	0.14
Black		Low WMC	0.33
	Stereotype-Inconsistent	High WMC	0.18
		Low WMC	0.49
	Stereotype-Consistent	High WMC	0.61
Asian		Low WMC	0.14
	Stereotype-Inconsistent	High WMC	0.1

Table 7. Source Monitoring Task Performance: Cued recall accuracy as a function of target race and WMC.

Target Race	WMC	Recognition Hits%
Black	Low WMC	28%
	High WMC	23%
Asian	Low WMC	28%
	High WMC	27%

Table 8. Source Monitoring Task Performance: Cued recall errors as a function of target race, word type, and WMC.

Target Race	Word Type	WMC	Recognition FAs %
		Low WMC	19%
	Stereotype-Consistent	High WMC	22%
Black		Low WMC	16%
	Stereotype-Inconsistent	High WMC	16%
		Low WMC	19%
	Stereotype-Consistent	High WMC	21%
Asian		Low WMC	18%
	Stereotype-Inconsistent	High WMC	18%

Table 9. Familiarity Task Performance: Accurate recognition (i.e., hits) as a function of target race, word type, and WMC.

Target Race	Word Type	WMC	Recognition Hits%
		Low WMC	87%
	Stereotype-Consistent	High WMC	86%
Black		Low WMC	87%
	Stereotype-Inconsistent	High WMC	89%
		Low WMC	90%
	Stereotype-Neutral	High WMC	92%
		Low WMC	89%
	Stereotype-Consistent	High WMC	94%
Asian		Low WMC	85%
	Stereotype-Inconsistent	High WMC	81%
		Low WMC	85%
	Stereotype-Neutral	High WMC	97%

Table 10. Familiarity Task Performance: Inaccurate recognition (i.e., false alarms) as a function of target race, word type, and WMC.

Target Race	Word Type	WMC	Recognition FAs %
		Low WMC	7%
	Stereotype-Consistent	High WMC	2%
Black		Low WMC	6%
	Stereotype-Inconsistent	High WMC	0%
		Low WMC	5%
	Stereotype-Neutral	High WMC	8%
		Low WMC	6%
	Stereotype-Consistent	High WMC	3%
Asian		Low WMC	3%
	Stereotype-Inconsistent	High WMC	0%
		Low WMC	1%
	Stereotype-Neutral	High WMC	0%

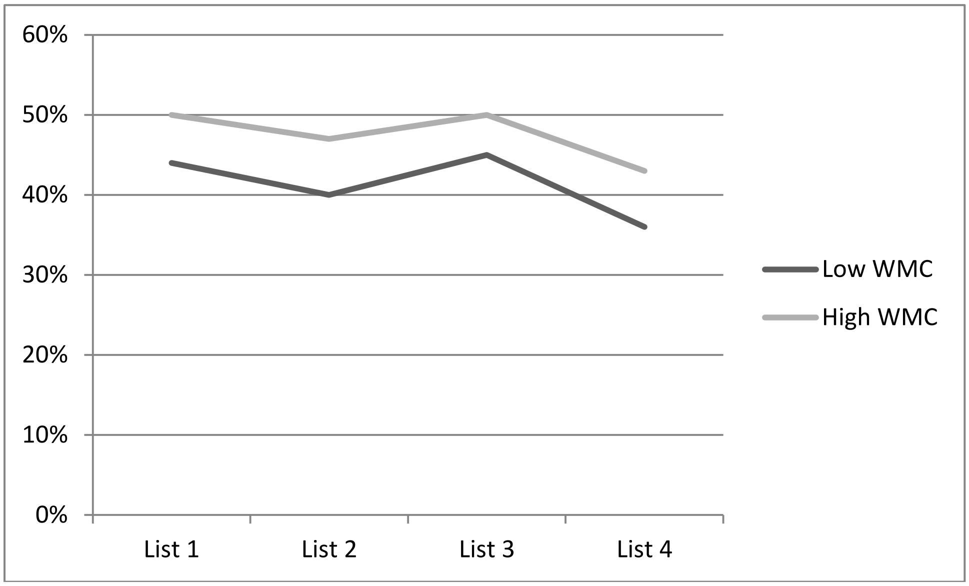

Figure 4. Retrieval Task Performance: Accurate recall by list as a function of WMC.

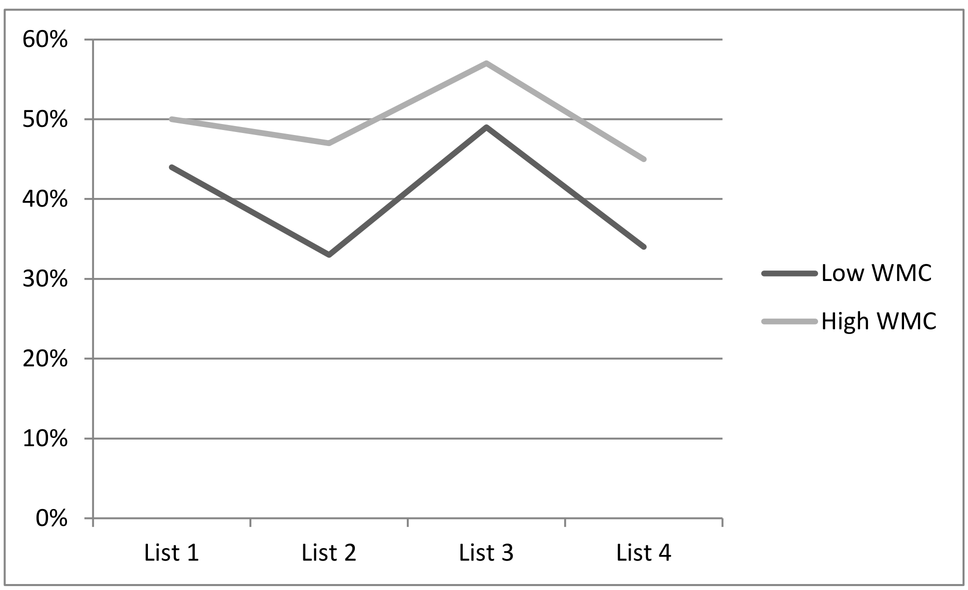

Figure 5. Retrieval Task Performance: Accurate recall by list for stereotype-consistent words as a function of WMC.

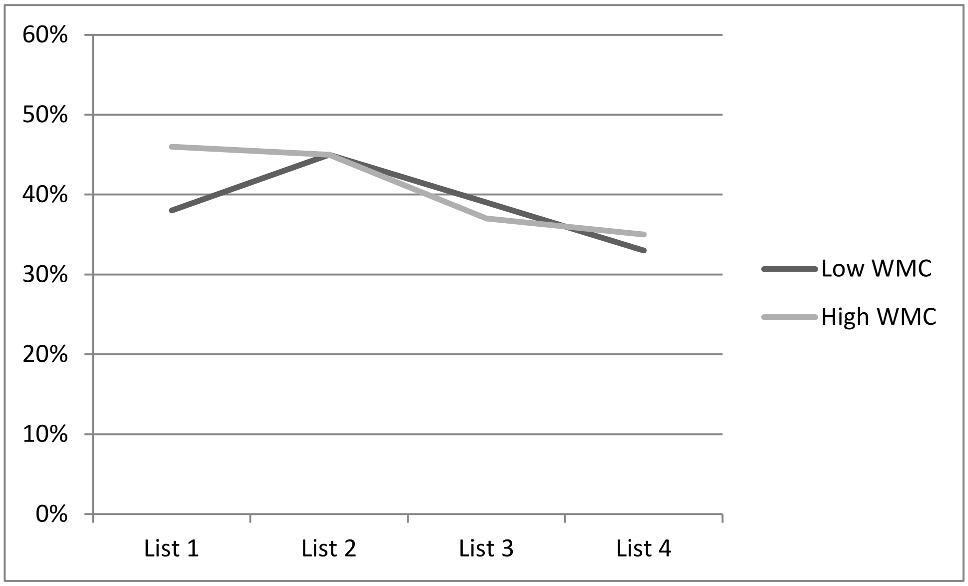

Figure 6. Retrieval Task Performance: Accurate recall by list for stereotype-inconsistent words as a function of WMC.

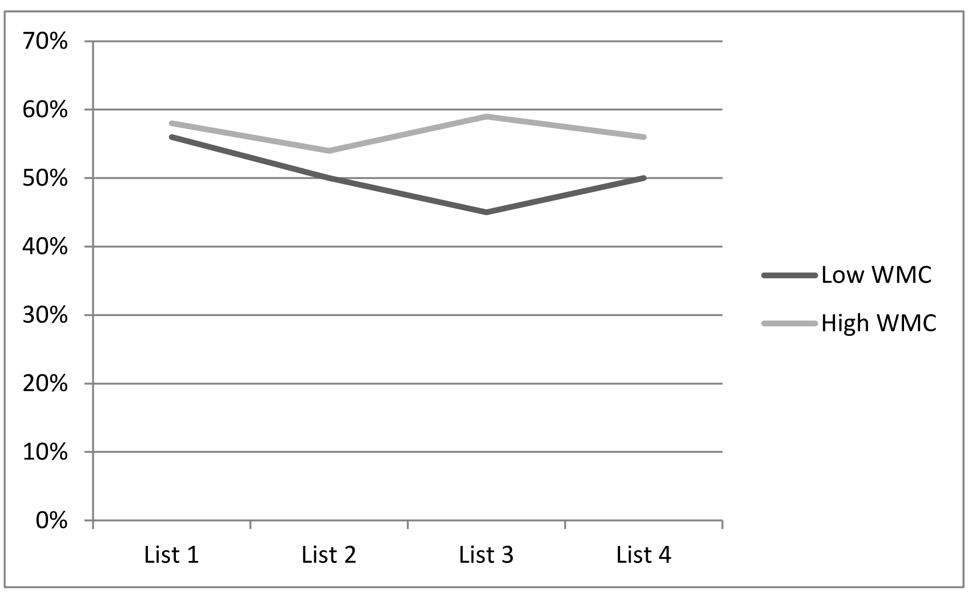

Figure 7. Retrieval Task Performance: Accurate recall by list for stereotype-neutral words as a function of WMC.