CALIFORNIA STATE UNIVERSITY SAN MARCOS

THESIS SIGNATURE PAGE

THESIS SUBMITTED IN PARTIAL FULLFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE

MASTER OF ARTS

IN

PSYCHOLOGY

THESIS TITLE: Individual Differences in Working Memory and Levels of
Processing

AUTHOR: David A. Illingworth

DATE OF SUCCESSFUL DEFENSE: August 9, 2011

THE THESIS HAS BEEN ACCEPTED BY THE THESIS COMMITTEE IN
PARTIAL FULLFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS IN PSYCHOLOGY.

Dr. Miriam Schustack
THESIS COMMITTEE CHAIR

SIGNATURE

DATE

Dr. Dustin Calvillo
THESIS COMMITTEE MEMBER

SIGNATURE

DATE

Dr. Marie Thomas
THESIS COMMITTEE MEMBER

SIGNATURE

DATE
Individual Differences in Working Memory Capacity and Levels of Processing

David A. Illingworth

California State University San Marcos
Abstract

Working memory (i.e., the capacity to maintain, manipulate, and retrieve multiple bundles of information) has been shown to be independent of, albeit correlated with, intelligence (g), and is variable across the population. As a result, the question has arisen as to what, if anything, accounts for differences in working memory capacity (WMC). Research addressing this question has had minimal success testing for differences in aptitude for encoding and retrieving between samples of individuals with high and low WMC. Prior research has been further limited by difficulties with empirically discriminating encoding and retrieval, as well as confusion about what factors related to encoding result in enhanced (or diminished) performance on measures of WMC. This study explored the relationship between encoding and WMC by controlling for depth of processing in a word recall paradigm (i.e., levels-of-processing task) that included a near-span length memory load in some conditions. The anticipated interaction between levels-of-processing and memory load was found, suggesting that deep processing is not susceptible to the effects of increased memory load in the same way memory loads affect lower domains of processing. All main effects were qualified by an interaction between WMC, memory load, and level-of-processing. Differences in patterns of performance between high and low WMC groups were only detectable at the deepest level of processing.

Keywords: working memory, levels-of-processing, attention, encoding
Individual Differences in Working Memory

Capacity and Levels of Processing

The development of the working memory construct has marked an important shift in the conceptualization of how people store, attend to, and manipulate information. Prior to Baddeley and Hitch (1974), the concept of a stagnant short-term memory store represented on-line cognitive processes. Alternatively, the multi-component model of working memory created by Baddeley and Hitch described an active process. It served to explain how memory was stored for brief periods of time, how sensory-specific memories were manipulated, and how the mind enacted control over these systems (Baddeley & Hitch, 1974). Recently, the creators of the model made an addition to their framework, proposing a method by which ideas were fused together and stored for brief periods of time and crediting working memory with focusing attention and inhibiting irrelevant information (Baddeley, 2003). Given the influence it is theorized to hold on various other cognitive phenomena, the most important property attributed to working memory is its finite capacity (Cowan, 2010).

Starting with the development of the reading-span task (Daneman & Carpenter, 1980), myriad measures have been developed to capture the limiting factor known as working memory capacity (WMC). The ability to scale individual capacities for online cognition has driven theory about the working memory construct, as the measures mentioned above have been used to compare working memory to other cognitive phenomena (e.g., intelligence). Since then, working memory has become a staple of cognitive science, and the focus of related research
has expanded to include searching for factors that may explain why WMC varies from one person to the next. Recent studies have revealed that attention, encoding, and retrieval are important variables in the study of individual differences in WMC (Unsworth & Spillers, 2010a; 2010b); however, an evaluation of the conclusions drawn by these studies suggests that limiting experimentation of individual differences to encoding phenomena may enhance understanding of individual differences given the links among the three variables mentioned. Encoding is entirely dependent on attending to a stimulus, and retrieval is likely dependent on encoding strategy for search cues. Controlling for attention strategies undertaken by subjects may be an important experimental manipulation for revealing how attention elicits variable performance on memory tasks. It may be the case that naturally occurring differences in the application of, or capacity for, attention is the encoding mechanism that explains variation in WMC across the population.

To better establish this hypothesis, the history and development of the individual differences approach to working memory will be presented as the foundation for the proposed research. Additionally, levels-of-processing will be discussed on its own, as well as within the context of primary and secondary memory systems. Working memory taxation will be described as a means to ensure that any effects discovered are a function of working memory.

Individual Differences in Working Memory

The study of individual differences in WMC dates back to Daneman and Carpenter (1980) who developed a reading-span task to assess the relationship
between WMC and reading comprehension. The reading-span task asked subjects to maintain, in memory, the final word from each of a series of sentences. The authors reasoned that the total number of words correctly recalled by subjects reflected the amount of information available to them for on-line cognition (i.e., working memory). This study inspired many researchers to develop additional measures of WMC and to expand upon the known influence of the working memory construct. This type of research remained the focus of the individual differences approach until questions arose as to whether or not WMC was an entity unto itself. Were measures of WMC tapping into an active memory system, or was the growing body of research on WMC reflecting the properties of some other cognitive process (e.g., general fluid intelligence)?

There have been many researchers who have theorized that working memory is the foundation of general intelligence (Conway et al., 2002; Deary, 2001; Engle et al., 1999); however, advancements in the understanding of the properties of working memory have put an end to that way of thinking. Attention studies, for example, have demonstrated that working memory contains a cognitive control factor that is not necessary for high intelligence. This difference between the two processes has been documented by neuro-imaging studies that have found distinctly different patterns of frontal lobe activity during measurement of WMC in comparison to activity during the measurement of general fluid intelligence (Conway, Kane, & Engle, 2003), such that regions of the brain believed to be responsible for cognitive control are only acting during working memory tasks. Conway, Cowan, and Bunting (2001) have
discovered that individuals with high WMC were better able to avoid making shadowing errors in a cocktail party paradigm than those with low WMC. Alternative dichotic listening tasks (i.e., capacity to detect target messages in an unattended ear) have also demonstrated enhanced performance from individuals with greater WMC (Colflesh & Conway, 2007). These studies illustrate that those with high WMC are capable of narrowing their attention when the goal is to ignore irrelevant information, and that they are able to expand their attention when they must juggle two tasks simultaneously.

The individual-differences approach to WMC has found evidence of an attention control factor in selective attention paradigms outside of dichotic listening tasks as well. Effects have been found in visual tasks such as saccade (i.e., eye movement) tasks (Kane at al., 2001) and the Stroop test (Kane & Engle, 2003). The most compelling finding unearthed by these studies suggests that individuals with high and low WMC perform equally well in conditions where attention demands are relatively low (e.g., prosaccade tasks, and matching word-color combinations in the Stroop test). The distinctions between high and low WMC individuals only appear once greater command of attention is demanded by the task (e.g., antisaccade tasks, and mismatched color-word combinations in the Stroop test).

Though there are substantial differences between general fluid intelligence and WMC, studies have shown that there exists a strong relationship between the two constructs (Engle et al., 1999). Recent studies have made attempts to discover what specific factors, capable of predicting performance on measures of WMC, explain the
correlations found with intelligence. A study comparing the operation-span task (a
dual-task measure of WMC) and the Raven Advanced Progressive Matrices (Raven)
measure of intelligence (two measures that typically load on separate factors despite
being highly correlated) discovered that their correlation was maintained across
various memory loads and levels of difficulty (Unsworth & Engle, 2005). Since the
expectation was that correlations would only be maintained within memory load
conditions (i.e., when memory was taxed to the same degree), the authors conclude
that their study is evidence that the relationship between WMC and intelligence goes
beyond general capacity.

A related idea supposes that WMC measures (e.g., operation-span task) differ
from general memory capacity measures (e.g., word span task) because they inhibit
rehearsal (due to dual-task designs) and, therefore, give a more precise measurement
of the capacity (i.e., the number of individual chunks) an individual has available for
maintaining information for the purpose of immediate cognition (Conway, Kane, &
Engle, 2003). Thus, the common factor between WMC and intelligence may be that
they are both limited by the true capacity an individual has for on-line cognition.
Added support for this hypothesis comes from a recent factor analysis of attention
control, secondary memory (i.e., the capacity to multitask), WMC, and intelligence,
which suggests that WMC continues to share unique variance with intelligence even
when properties of working memory (i.e., attention control and secondary memory)
are controlled for (Unsworth & Spillers, 2010a).
Recently, research has begun to explore the relationship between individual differences in WMC and retrieval. Earlier studies had established grounds for studying this phenomenon, as they had discovered that there were significant differences between individuals with high and low WMC in their ability to perform fact-retrieving tasks (Cantor & Engle, 1993; Conway & Engle, 1994). Unsworth (2007) has found that in a world-list recall task, individuals with low WMC were more likely than individuals with high WMC to experience interference from previous word-lists and reported fewer words in total. Additionally, individuals with low WMC recalled words at a slower rate than those with high WMC. Subsequent studies have been successful at replicating this effect; however, expected differences in their responses to specific cues for retrieval were not found (Unsworth, 2009). That is, when cues designed to assist subjects in searching a smaller set of information are used, individuals with low and high WMC experience the same rate of enhanced recall.

Stemming from the initial interest of discovering the realm of influence for working memory is the question as to what, if anything, accounts for individual differences in WMC. Attention remains a fixture in the body of literature, as there exists the idea that individuals vary in their capacity to adjust their attentional focus (Cowan et al., 2005). However, a separate body of literature has begun to address this issue in terms of individual differences in encoding and retrieval. This has arisen in response to the research mentioned previously demonstrating that individuals with high and low WMC search for information in different ways. Moreover, studies have
suggested that attention may be involved in this relationship, as it may enhance the quality of encoding (Brewer et al., 2010). Additional evidence is provided by change detection studies that have suggested that a lack of selection (i.e., the assignment of attention) during encoding accounts for differences in performance on measures of visual WMC (Cusack, Veldsman, & Mitchell, 2009).

The only direct empirical test of encoding and retrieval as explanations for individual differences in WMC produced mixed results (Unsworth & Spillers, 2010b). In this study, the effects of encoding and retrieval were differentiated by manipulating whether the memory task (i.e., word-span task) was explicit or implicit. Individuals with high WMC outperformed those with low WMC; this was especially true in the explicit memory task condition. The authors concluded that people with high WMC were better equipped for using advanced encoding strategies and using cues for retrieval. These assumptions appear overreaching when one considers whether or not individuals with high WMC are likely to perform better than those with low WMC on any task. Meta-analysis has demonstrated that WMC is a predictor for performance on many cognitive measures (Ackerman, Beier, & Boyle, 2005). Therefore, the finding that high WMC outperformed low WMC individuals is not evidence enough for any specific mechanism. Additionally, no manipulation was introduced to control for any encoding that may have occurred in the implicit task. This study does, however, add to the understanding of the interplay of these variables, as they discovered that there was a substantially greater difference in performance during the explicit task than the implicit task. Enhanced performance during explicit
tasks suggests that encoding may better explain differences in WMC than retrieval and could be a more fruitful pursuit for future research.

**Levels of Processing**

Craik and Lockhart (1972) suggested that people processed information at different levels, and theorized that deeper levels-of-processing should result in improved recall of a stimulus in comparison to shallow levels-of-processing. This idea gained empirical support when Craik and Tulving (1972) demonstrated that individuals were better able to recall words when they were processed at the semantic level (i.e., categorizing the word) than when they were processed at auditory (i.e., rhyming the word) or structural levels (i.e., assessing the case of the word). Levels-of-processing remains a robust effect, commonly used to assess memory processing of abnormal populations (Bonner-Jackson, Csermansky, & Barch, 2007; Lespinet-Najib et al., 2004; Paul, Elvevag, Bokat, Weinberger, & Goldberg, 2005, Ragland et al., 2003; Whitehouse, Maybery, & Durkin, 2007) and aging populations (Dulaney, Marks, & Link, 2004; Froger, Taconnat, Landre, Beigneux, & Isingrini, 2009; Ramponi, Richardson-Klavehn, & Gardiner, 2004; Sauzeon, N’kaoua, Lespinet, Guillen, & Claverie, 2000).

Despite its continued use, the theory was subject to multiple criticisms soon after its initial publication. These were primarily due to uncertainty about the rigidity of levels-of-processing, the possibility that structural prompts (e.g., Is the word in lower case?) produced deeper levels of processing than what was expected (Nelson, Walling, & McEvoy, 1979), and difficulties in applying the concept outside of verbal
paradigms (Baddeley & Woodhead, 1982). Due to the vague definition of depth and the lack of an objective hierarchy for domains of processing, the levels-of-processing paradigm was also criticized as relying on circular logic (Lockhart & Craik, 1990). Deeper processing was assumed to have occurred when recall was enhanced, thus experimental manipulations that elicited that enhanced performance were labeled as deep processing. Lockhart and Craik (1990) responded to these criticisms by first stating that their understanding of processing was not fully encapsulated by the three manipulations introduced by their earliest studies. Domains of processing are likely to produce a messy hierarchy and would likely require performance based comparisons, as their early work refutes the use of independent measures of depth such as the amount of time spent processing a stimulus.

Contemporary versions of the levels-of-processing tasks have adjusted for these early criticisms by applying the concept of distinctiveness (Gallo, Meadow, Johnson, & Foster, 2008). It is assumed that subjects will process a stimulus at a semantic level regardless of the condition, unless a manipulation designed to inhibit such processing is worked into the paradigm. It is proposed that drawing a subject’s attention to specific details of a word, so long as the entire word is not processed, will result in shallow processing, as doing so causes words to become less distinctive. Subjects are asked to count letters (e.g., number of vowels) for lower level processing, and asked to judge the pleasantness of a word for deeper processing (Hyde & Jenkins, 1969).
Additional criticisms aimed at the levels-of-processing paradigm concern the utility of trace memory theories in general (i.e., that memory must incorporate an encoding phase, retention, and retrieval; Clifford, 2004; Watkins, 1990), and the fact that the levels-of-processing framework seems to focus on the encoding phase of memory, ignoring the other processes (Kolers, 1975; Kolers & Ostry, 1974). For the purposes of this study, levels-of-processing is being used as a tool to study a memory process. That is, the specific theories associated with processing are not inherently important for the consideration of the current study. Even the critics of levels-of-processing suggest that the paradigm is related to encoding, and that is its only role in this study.

Although theoretical advances have not been pursued within the levels-of-processing paradigm (Roediger, Gallo, & Geraci, 2002), support for the framework has grown due to assessment of the paradigm in neuro-imaging studies. Nyberg (2002) tested the hypotheses that sensory information about a stimulus (i.e., perceptual brain activation) does not fade away immediately, and that brain regions associated with deeper levels of processing should correspond with enhanced recall performance. Both of these predictions were supported, adding credibility to levels-of-processing (Nyberg, 2002). Working memory has been implicated in neuro-imaging studies of levels-of-processing, as tasks associated with each level (i.e., structural, auditory, and semantic) have elicited activation in the same regions of the frontal lobe (Innocenti et al., 2001).
When discussing WMC, secondary memory and attention are often reported as possible factors that explain differences in WMC, and it has been noted frequently that WMC is highly correlated with measures of intelligence (Unsworth & Spillers, 2010a). Despite the minimal advances in theory, researchers have expanded upon the concept enough to make similar claims about levels-of-processing concerning cognitive ability, attention, and secondary memory. Schweizer (1996) found that performance on a deep level of encoding task correlated with, and best predicted, enhanced performance on cognitive ability tasks (i.e., the Raven intelligence measure and a reasoning task). WMC (i.e., operation-span performance) has already been discussed previously as correlating with Raven’s matrices. Other research has found that levels-of-processing and attention work in the same way towards enhancing recall on explicit memory tasks (Bentin, Moscovitch, & Nirhod, 1998). These studies provide grounds for expecting levels-of-processing to account for differences in WMC.

Only two direct tests of the relationship between levels-of-processing and WMC have been completed, and these had mixed results. One study found that developing a new measure of WMC that requires subjects to process words at various levels (i.e., visual, phonological, and semantic) does not produce enhanced performance at deeper levels-of-processing (Rose, Myerson, Roediger, & Hale, 2010). However, due to the integration of the WMC measure and the levels-of-processing manipulation, it is difficult to conclude that there is no relationship between the constructs, as no change would be expected if levels-of-processing is
inherently worked into measures of WMC. On the other hand, McCabe (2008) found that deeper processing improved delayed recall for items encoded during the operation-span task in comparison to a word-span task.

**Working Memory Load**

As mentioned previously, WMC is capable of predicting how individuals will perform on a variety of cognitive tasks (Ackerman, Beier, & Boyle, 2005). Thus, any conclusions drawn from data assessing individual differences in WMC must be careful to consider this fact, avoiding substantial claims and causal relationships when general cognitive ability is capable of explaining results. Use of a memory load is, therefore, a crucial manipulation for ensuring that any effects discovered in a study are the result of other controlled variables in any experimental paradigm.

Dual task studies have become increasingly popular in the working memory literature and have used several different paradigms to assess the overlap or conflict between cognitive functions. These tasks are believed to weigh on working memory and its components, further limiting its capacity to function. Complicating the matter of inducing a working memory load is the fact that disagreement exists over the structural make-up of working memory. While some theorists believe that working memory is a multi-component structure (Baddeley, 2003; Baddeley & Hitch, 1994), there are others who believe that working memory itself is but a small component in a larger memory construct (Ericsson & Kintsch, 1995). Thus, two questions arise concerning working memory and dual tasks: Does empirical evidence support a
multi-dimensional working memory model and, if so, do different types of dual tasks affect some working memory components and not others?

Baddeley and Hitch’s (1994) theoretical framework for a multi-component working memory consists of three different systems. These are the visuospatial sketchpad (a short term storage system for visual and spatial information), the phonological loop (a short term memory store for speech information), and the episodic buffer (a short term memory store for episodic memory; Baddeley, 2003); the central executive (an attention focusing mechanism) organizes and ties all three systems together. Studies that support the Baddeley and Hitch model have shown that processes exist that are responsible for maintaining different kinds of stimuli for brief periods of time. The visuospatial sketchpad, for example, has primarily been established by studying the maintenance of visual information in abnormal cases. In a study conducted by Nissen et al. (1985), patients with Alzheimer’s disease and “normal” individuals were presented with bars that differed from their backgrounds in varying levels of luminosity, and were asked to report whether or not they could detect the bars. Results demonstrated that the degenerative properties of Alzheimer’s (i.e., depletions in memory capacity) also had an effect on retaining visual information over brief periods of time, as Alzheimer’s patients performed poorer on the task at all levels of luminosity. Although examining the visuospatial sketchpad is much more difficult than examining the phonological loop, several studies with similar designs have produced parallel patterns of data.
Unlike the visuospatial sketchpad, the phonological loop has been studied at length. Several research designs have been able to discover stimuli for which this element is sensitive. Some of the most compelling evidence for this mechanism has come from studying word length and how working memory facilitates the encoding of words. Caplan, Rochon, and Waters (1992) conducted a study where words of different syllabic length were presented to subjects to determine if there were differences in the amount of working memory span required to process and maintain long words in comparison to short words. Words with more syllables required more cognitive effort than words with fewer syllables. Furthermore, when subjects were allowed to collapse syllables, so that longer words could be pronounced in fewer syllables (e.g., suppose to s’pose), the necessary working memory span decreased. Results demonstrate that when longer words are shortened into fewer syllables, their tax on working memory is significantly lessened. The relationship between phonological word length and working memory supports the claim that working memory has a short term phonological store that can become taxed as it takes on more information.

The central executive is the least studied, and least understood component of the Baddeley and Hitch (1994) working memory model. Conceptually, it is regarded as the binding force of working memory as a whole. Beyond acting as the executive branch of working memory, one may speculate that as the organizing entity, the central executive is responsible for reducing interference from irrelevant information, focusing attention, and other organizational tasks. Given the nature of these
behaviors, however, it is difficult to see how these relationships can be studied empirically. Baddeley and Hitch suggest that dual tasks may be a means for measuring the efficiency of the central executive, as it should be responsible for controlling what is held in focus within the mind.

One reason why the central executive may be so poorly understood is the misattribution of properties to the component that actually belong to another. Baddeley (2001) has recently proposed a novel component to his model called the episodic buffer. As a result, the central executive has been reduced to an attention focusing module, as the episodic buffer takes on the role of binding the model together. Data has shown that long-term memory may play a role in sentence span (i.e., the number of words an individual can recall in prose) that is absent in typical word span measures (Baddeley, Vallar, & Wilson, 1987). Therefore, it has been proposed that the episodic buffer also serves as the mechanism by which long-term knowledge enters working memory. Neuro-imaging studies support the existence of such a mechanism (Prabhakaran, Narayanan, Zhao, & Gabrielli, 2000).

The question of whether or not working memory is resistant to retaining an information load on one element of the system while performing a task with another remains. Studies have used dual tasks that are dependent on the same element within working memory and, while they demonstrate that WMC is subject to depletion under dual task conditions, they fall short of testing a generalized capacity theory. For example, Han and Kim (2008) measured the extent to which a visuospatial memory load interfered with a visuospatial task. Under dual task conditions, subjects were
shown an array of square markers and asked to maintain their positions in short term memory. The focal task measured if subjects reacted faster to a target on a screen after they had been visually cued to where it would appear than when the target appeared without warning. Findings demonstrated that a visuospatial load on working memory resulted in a reduction in accuracy in the visual acuity task and increased the reaction time to complete the task with or without a visual cue.

In order to test the resilience of working memory against general taxation, dual tasks with visual and verbal components must be adopted. In a study attempting to measure the visual binding abilities of a proposed new element to their working memory theory, Allen, Baddeley, and Hitch (2006) induced a verbal working memory load with the goal of interfering with the fusion of visual details. The focal task required that subjects judge whether or not a target object, an object of a specific color, or a combination of both the color and shape was present during an initial presentation of objects. Two different dual task conditions were used in this study: one required that subjects count backwards starting at a randomly assigned 3-digit number during the encoding of the initial presentation of objects, while the other required subjects to maintain and recall the digits one through six which were presented on a computer screen in random order. Neither of these tasks can be considered to be visual in nature, and yet both were shown to cause a significant reduction in accuracy when recalling the shapes and colors presented.

Other theorists have been more direct in their assertion that working memory can be depleted regardless of the type of secondary task that is presented to a subject.
Conway and Engle (1996) tested three different theories, which predicted different tradeoffs in response to dual task conditions. Their paradigm consisted of finding correlations between an individual’s success on a secondary task and scores on their verbal SATs (reading comprehension). Results demonstrated that, despite the different components of working memory and the distinctions across processing types, the limitations of working memory could best be described using a general capacity theory. The data suggest that when a dual task is used in an experimental paradigm, WMC, in its entirety, is depleted.

The Present Study

The present study builds upon the research discussed previously by testing a novel explanation for individual differences in WMC. Though encoding and retrieval have been demonstrated to be important factors in WMC (Unsworth & Spillers, 2010b), no one has attempted to discover what specific encoding behaviors result in enhanced capacity for online cognition. Patterns of free recall performance for individuals with low and high WMC across various memory load conditions in a levels-of-processing paradigm were analyzed to test the hypothesis that differences in WMC are the product of naturally occurring differences in processing.

The goal of this study was to answer two questions. First, is there a difference between individuals with high and low WMC in terms of their performance on a levels-of-processing task? That is, do individuals who score at the extreme ends of a working memory measure respond differently to deeper levels-of-processing? It was hypothesized that 1A: Individuals with high WMC will perform better on all recall
tasks regardless of the depth of processing condition, and 1B: The effects of deeper processing will differ between the two focal groups such that individuals with low WMC will have a markedly greater increase in words recalled than those with high WMC as the required depth of processing increases.

This study also addresses whether or not differences elicited by levels-of-processing are a function of WMC. By adding two conditions that include a secondary task (i.e., a near-span length memory task), this study will empirically test whether limiting an individual’s capacity to attend to and manage information inhibits performance on a levels-of-processing task. It is expected that 2A: All subjects will demonstrate an increased ability to recall words as processing deepens, while 2B: Individuals with high WMC will outscore those with low WMC regardless of the memory load; and 2C: There will be an interaction between the level at which stimuli are being processed and the memory load.

**Method**

**Participants**

One-hundred and fifty-eight students were recruited from the department of psychology’s Human Participant Pool at California State University San Marcos. Undergraduate students enrolled in lower division psychology courses received partial credit toward course requirements for their participation. Of the original 158 subjects, 38 were excluded from analysis due to computer failure (\(N = 25\)), incomplete data (\(N = 9\)), or researcher error (\(N = 4\)). The remaining subjects (\(N =\))
120) had a mean age of 20.13 ($SD = 2.09$), ranging from 18 years of age to 28.

**Design**

The design was a 4 (working memory capacity; low, below average, above average, and high) X 3 (levels-of-processing; shallow, moderate, and deep) X 3 (memory load; none, low, and high) mixed model factorial, with memory load as a blocked, within-subject variable. Levels-of-processing served as a second within-subject variable and WMC was a between-group variable.

The WMC groups were established using a working memory measure known as the operation-span task. It should be noted that although all four WMC groups were included in the initial analyses, only differences between the extreme quartiles are of interest for assessing differences between low and high WMC. This plan of analysis avoids issues that arise when continuous variables are arbitrarily grouped and creates a sizable gap between the groups of interest.

**Measures**

**Levels-of-processing.** The levels-of-processing task was completed over three blocks. Each block included one level of the working memory load but all three levels-of-processing tasks. Blocks consisted of 60 trials (i.e., question-word pairs), wherein 20 trials used the shallow processing task, 20 trials used the moderate processing task, and 20 trials used the deep processing task. Previous research has shown that attending to structural features of a word can alter performance long after presentation of the stimulus (Conway & Gathercole, 1987; Lockhart & Craik, 1990), thus eliciting memory effects that do not reflect shallow processing. In this study,
shallow processing was elicited by prompting subjects to count the number of vowels in the stimulus (e.g., Vowels?), and report that number (e.g., 2) in a box that appeared on the screen following the presentation of the stimulus (e.g., Wood). For moderate processing trials, subjects were asked to read the stimulus without a response of any kind (e.g., Read.) Deep processing trials asked subjects to rate the pleasantness of the meaning of a word on a rating scale ranging from 1 (very unpleasant) to 5 (very pleasant) (e.g., Pleasant?). Subject responses (e.g., 5) were reported in a box that appeared on the screen after the presentation of the stimulus (e.g., Vanilla). Detailed instructions and practice sessions were presented prior to the experimental conditions to ensure that subjects understood the prompts that appeared on the screen. Each question was visible for 1000 ms prior to the presentation of the word. Words were presented for 3000 ms regardless of the question type or how quickly subjects responded to the stimulus.

Performance on the levels-of-processing task was measured by subject performance on a three minute long free recall task following each set of words. The recall was scored by counting the number of responses that matched words that were presented during the block. Words were accepted so long as they were recognizable to the judge in either singular or plural form. Scoring was completed by two judges who were blind to condition. When discrepancies appeared in the scoring of the data, a third judge scored the responses.

The effects of levels-of-processing are typically measured using a recognition task (Craik, 1977; Moeser, 1983) in addition to free recall, however, recent studies
have demonstrated that performance on both free recall and the recognition task are the result of the distinctiveness heuristic produced by levels-of-processing (Gallo, Meadow, Johnson, & Foster, 2008). Therefore, to reduce subject fatigue, free recall was used as the only measure of performance.

**Reliability of Scoring.** Reliability was tested by calculating the frequency with which the two judges agreed and the correlation of their scores. In this data set, the judges agreed on the number of words recalled for 60.60% \( (n = 218/360) \) of the blocks. The most common reason for disagreement was uncertainty allocating credit for a word when the subject demonstrated poor spelling. The other reason for disagreement was failure to find a reported word in the correct stimuli set. The percentage of agreement jumps up to 87.20% \( (n = 314/360) \) when the difference between judges was equal to zero or one word, and 93% \( (n = 336/360) \) when the difference was equal to two or fewer words. The correlation between the two judges, on the other hand, was very high, \( r = .93, p < .001 \).

**Working memory load.** The secondary task to the levels-of-processing task was a near-span length verbal memory task. This task asked subjects to maintain six digits (presented in pairs) in their short term memory, while performing the primary task. Successful completion of this task requires that subjects recall the digits in the order in which they were presented. Three memory load conditions were used in this study. In the “no load” condition, there were no digits presented. In the “low load” condition, digits were presented in cardinal order (e.g., 2-3-4-5-6-7). In the “high load” condition, six random digits were presented to subjects (e.g., 7-2-4-1-5-9).
Memory loads were induced prior to the presentation of word sets in the levels-of-
processing task. The memory load was made prominent by prompting subjects to
report the digits prior to the 3 minute long free recall that followed the levels-of-
processing task. Practice trials demonstrated to the subjects that there was no further
need to maintain the digits in memory after they had been reported.

The efficacy of secondary tasks is measured in two ways: do subjects perform
well enough on the task to suggest that cognitive resources were spent completing it,
and to what extent does it interfere with the focal task? Previous studies have
demonstrated that subjects generally perform equally well (approximately 80%
success-rate) on the near-span length verbal memory task regardless of the nature of
the focal task (Allen, Baddeley, & Hitch, 2006). This paradigm has been
successfully used to interfere with cognitive functions such as image binding (Allen,
Baddeley, & Hitch, 2006), delayed discounting (Hinson, Jameson, & Whitney, 2003),
and risky decision making (Whitney, Rinehard, & Hinson, 2008). The script used for
this study failed to store responses to this portion of the study. Thus, it is impossible
to report on subject performance on the secondary task.

**Working memory capacity.** The operation-span task is widely accepted as a
reliable and successful measure of WMC. Additionally, it has become the most
common measure of WMC in individual differences research. Developed as a means
to test the relationship between WMC and upper-level thinking and comprehension
(Turner & Engle, 1989), the o-span task has proven to be a valid measure of memory
span (Conway & Engle, 1996). Having examined the test-retest reliability of the
operation-span task, Klein and Fiss (1999) determined that the task is reliable at both
the 2-3 week interval ($r = .667$) and 6-7 week interval ($r = .812$). For the purposes of
this study, the automated o-span task was used. This adaptation of the original task
has been shown to have similar psychometric properties (Unsworth, Heitz, Schrock,
& Engle, 2005).

This task required that subjects maintain a series of letters in mind while
making true/false judgments about mathematical operations. Operations and letters
were presented in pairs wherein the operation is presented first (e.g., $8/4-1=, 2*3+3=)$.
Subjects are told to solve the problem in their mind before clicking to the next screen
where they are presented with a solution to the operation that may or may not be
correct (e.g., 1, 12). After judging the presented solution as either true or false, a
single letter was presented for one second. Each block consisted of 3 to 7 trials
before the subjects were prompted to recall the letters in the order that they were
presented. Each subject saw a total of 15 blocks, and received a point for each letter
correctly reported when every letter within the block was correctly recalled. For
example, had a subject correctly reported 6 letters in a block that contained 7
operation-letter pairs, the subject would have received 0 points. Had the subject
correctly reported all 7 letters, they would have received 7 points. The highest
possible score was 75. The operation-span scores of the final sample ($M = 36.38, SD$
$= 17.22$) resembled those of previous samples (Illingworth & Schustack, 2010;
Unsworth, Heitz, Schrock, & Engle, 2005). For a comparison with other samples, see
Table 1. To see how this measure was used to create the WMC groups, see Table 2.
Materials

The stimuli for the levels-of-processing task were drawn from the online Medical Research Council (MRC) database (Coltheart, 1981; Wilson, 1988). All stimuli were controlled for word length (5 to 7 letters), written frequency (greater than 1), and familiarity, concreteness, and imagability ratings (500-700 on 100-700 scales). Given these criteria, the MRC database produced 380 words. These stimuli were randomly assigned to 6 sets of 60 words; the remaining 20 words were assigned to a practice trial list.

All stimuli and instructions were presented in Arial (size 16) font on a 19-inch Dell brand color monitor using the E-prime 2.0 software package. All key-pressed responses, time latency, and response accuracy (o-span) were recorded using this program.

Procedure

After giving informed consent and reporting basic demographic information (e.g., age, gender, ethnicity), subjects completed a series of practice trials that exposed them to the near-span length verbal memory task (i.e., the memory load) and the levels-of-processing task. Then, subjects completed three blocks of the levels-of-processing task. For each block, one of the six stimuli sets was randomly selected (without replacement). The order of the memory load conditions was randomized for each subject, as was the order in which stimuli were presented within each block. The order of events in the study is shown in Figure 1.
At the start of the study’s second phase, subjects were presented with a series of practice trials on the operation-span task. After final instructions were given to the subject, they completed 15 blocks that ranged from 3 to 7 trials each, totaling 75 trials overall.

**Results**

**Main Analysis**

A 4 (WMC) x 3 (Levels-of-processing) x 3 (memory load) mixed model ANOVA was run to test for differences in the number of words recalled. The assumption of sphericity was violated (levels-of-processing: $\chi^2 = 44.435, p < .001$; memory load by levels-of-processing interaction: $\chi^2 = 20.62, p < .05$); therefore, the following analyses were calculated using a Greenhouse-Geisser correction (levels-of-processing: $\varepsilon = .76$; memory load by levels-of-processing interaction: $\varepsilon = .92$).

Normality was also violated ($p < .05$), however, attempts to normalize the data failed due to the frequency with which subjects failed to report any words for specific conditions within a block of trials. This may be evidence of the difficult nature of maintaining 60 stimuli in memory for brief periods of time.

**Main Effects.** The main effect for memory load was found to be statistically significant ($F(2,232) = 5.44, p < .01, R^2 = .05$), as was the main effect for levels-of-processing ($F(1.51,175.68) = 37.62, p = < .001, R^2 = .25$). The between-subject variable (WMC group) did not cause statistically significant differences in words recalled ($F(3,116) = 0.99, p = .40$). Additionally, neither memory load ($F(6,232) = \ldots$)
1.45, \( p = .20 \) nor levels-of-processing \((F(4.54,175.68) = 1.18, \ p = .32)\) had a statistically significant interaction with WMC group.

**Two-way interaction.** These main effects were qualified by a marginally significant, two-way memory load by levels-of-processing interaction \( (F(3.68,426.60) = 2.00, \ p = .09, R^2 = .02)\), indicating that memory loads did not have the same effect on word recall at all levels of processing (Figure 2). Follow-up analyses reveal that the no memory load condition \((M = 2.57, \ SE = .20)\) did not differ significantly from the low memory load condition \((M = 2.38, \ SE = .21; \ t(119) = 0.83, \ p > .05)\) under shallow processing, but was statistically different from the high memory load condition \((M = 2.13, \ SE = .19; \ t(119) = 1.99, \ p < .05, \ d = .36)\). At the moderate processing level, no memory load \((M = 4.03, \ SE = .22)\) did not differ from low memory load \((M = 3.85, \ SE = .23; \ t(119) = .79)\), but was statistically different from high memory load \((M = 3.23, \ SE = .23; \ t(119) = 3.60, \ p < .05, \ d = .66)\). Finally, at the deep processing level, no memory load \((M = 4.26, \ SE = .25)\) was not statistically different from low memory load \((M = 4.33, \ SE = .25; \ t(119) = .30, \ p > .05)\) or high memory load \((M = 4.28, \ SE = .25; \ t(119) = .08, \ p > .05)\).

**Three-way Interaction.** The previously mentioned effects were also qualified by a three-way WMC group by memory load by levels-of-processing interaction \((F(11.03,426.60) = 1.79, \ p = .05, R^2 = .04)\). For the low WMC group, at the shallowest level of processing, no memory load \((M = 2.30, \ SE = .39)\) was not statistically different from low memory load \((M = 2.03, \ SE = .42; \ t(119) = 1.20, \ p > .05)\), but no memory load did differ significantly from high memory load \((M = 1.63, \ SE = .25; \ t(119) = 4.31, \ p < .05, \ d = .66)\).
$SE = .37; t(119) = 3.00, p < .05, d = .14)$. The low WMC group (Figure 3) showed no difference between the no memory load condition ($M = 4.0, SE = .44$) and the low memory load condition ($M = 3.57, SE = .46; t(119) = .43, p > .05$) at the moderate level of processing, but recalled significantly fewer words at the high memory load condition ($M = 3.20, SE = .47$) than they did during the no memory load condition ($t(119) = 3.60, p < .05, d = .66$). At the deep level of processing, the number of words recalled during the no memory load condition ($M = 4.20, SE = .50$) was not significantly different from either the low memory load condition ($M = 3.83, SE = .49; t(119) = 1.65, p > .05$) or the high memory load condition ($M = 4.27, SE = .49; t(119) = .30, p > .05$).

The high WMC group (Figure 4), at the shallow level of processing, performed no differently with no memory load ($M = 2.8, SE = .39$) than they did with a low memory load ($M = 2.43, SE = .42; t(119) = 1.65, p > .05$). Performance with no memory load was, however, statistically different from performance with a high memory load ($M = 1.93, SE = .37$) at the same level of processing ($t(119) = 3.90, p < .05, d = .71$). At a moderate level of processing, high WMC individuals did not perform differently with no memory load ($M = 4.53, SE = .44$) than they did with a low memory load ($M = 4.37, SE = .46; t(119) = .75, p > .05$), but when no memory load performance was compared with high memory load performance ($M = 4.07, SE = .47$) a statistically significant difference was found ($t(119) = 2.09, p < .05, d = .38$). When processing words at the deepest level, high WMC individuals recalled significantly fewer words with no memory load ($M = 3.87, SE = .50$) than they did
with a low memory load ($M = 4.90, SE = .49; t(119) = 4.64, p < .05, d = .85$).

However, there was no significant difference between their performances with no memory load and high memory load ($M = 4.30, SE = .49; t(119) = 1.94, p > .05$).

**Additional Analyses**

**Gender and Ethnicity.** Additional analyses were run to test for the influence of other variables on the number of words recalled by the subjects. Using total words recalled and operation-span scores (absolute) as dependent variables, one-way ANOVAs found no effects for gender (total words: $F(1,118) = 1.62, p > .05$; o-span: $F(1,118) = 1.10, p > .05$) or ethnicity (total words: $F(5,114) = .61, p > .05$; o-span: $F(5,114) = .19, p > .05$).

**Experimenter.** One-way ANOVAs produced split results for experimenter, as no effect was found on the total number of words recalled ($F(2,117) = 1.37, p > .05$) but a statistically relevant effect was found on o-span scores ($F(2,117) = 3.25, p < .05$). Follow-up analyses reveal that the experimenter with the highest average ($M = 42.84, SE = 3.04$) elicited significantly higher operation-span scores than the experimenter with the second highest average ($M = 35.08, SE = 2.32; t(38) = 2.05, p < .05, d = .66$). The difference between the second highest average and the lowest average ($M = 32.75, SE = 2.82$) was not statistically significant ($t(42) = .62, p > .05$). Neither of the dependent variables (i.e., total words and o-span) had statistically significant correlations with age (total words: $r = -.15, p > .05$; o-span: $r = .12, p > .05$) or hour of the day (total words: $r = -.06, p > .05$; o-span: $r = .09, p > .05$).
**Fatigue.** A mixed model ANOVA was calculated to test for fatigue effects on words recalled (i.e., did the total number of word correctly reported change from the first block of trials to the final block of trials?). No such effect was found ($F(2,232) = 0.23, p > .05$), nor was there an interaction with WMC groups ($F(6,232) = .48, p > .05$). A second mixed model ANOVA was calculated to assess the effect of the number of blocks seen by the subject on interference (i.e., words reported that had appeared among the stimuli presented in a prior block). The assumptions of normality ($p < .05$) and sphericity ($\chi^2 = 6.72, p < .05$) were violated in this analysis. A Greenhouse-Geisser correction was used in the calculation of the following statistics ($\varepsilon = .95$). Interference was found to be significantly altered by the number of blocks seen by the subject ($F(1.89,219.54) = 19.15, p < .001, R^2 = .14$). A trend analysis of this effect reveals a statistically significant linear increase (Figure 5) in interference as the number of blocks seen by the subject increases ($F(1,116) = 32.33, p < .001, R^2 = .22$).

**Stimuli Sets.** Data were also collected concerning subject performance in response to the six different sets of stimuli used in this study. However, these data were dependent by nature and incomplete, as each subject only saw three of the six sets. Thus, no statistical analysis can be conducted to assess the effect of stimuli set on the number of words recalled. Descriptive statistics for each stimuli set can be found in Table 3.

**Discussion**
The present study sought to test a novel hypothesis concerning the way encoding strategies may influence WMC. As predicted, individuals with high WMC were found to perform better on all recall tasks regardless of the depth of processing condition (Hypothesis 1A). However, this phenomenon was not statistically significant. This may be the result of averaging performance across all levels of memory load and processing. The effects of deeper processing did not differ between the two WMC groups (Hypothesis 1B). Prior to the experiment, it was believed that individuals with low WMC might have had a markedly greater increase in performance as the required depth of processing increased, in comparison to those with high WMC. Visual inspection of the results (see Figure 6) may suggest that individuals with high WMC plateau as soon as a moderate processing strategy is undertaken, while those with low WMC improve in the number of words recalled gradually as processing deepens. Unfortunately, this pattern is not statistically reliable.

The design of the levels-of-processing paradigm for this study was influenced by Gallo et al. (1980), who strongly advocated inhibited processing of words under shallow conditions. The desired effect was found, as both WMC groups performed poorly when they were asked to count the number of vowels in a word. However, it must be taken into consideration that such inhibition may not result in any processing of the word at all. What the subjects do attend to—vowels—are poor cues for retrieval, as they are common elements in every word in existence. Thus, to better apprehend how depth of processing interacts with WMC, future studies may want to
incorporate more intervals in the processing spectrum. This study is limited in its ability to draw finite conclusions about this relationship given only two data points where processing of stimuli is known to have occurred were available for each WMC group.

The second series of hypotheses in this study were supported, as all subjects did enjoy increased ability to recall words as processing deepened (Hypothesis 2A), while individuals with high WMC outscored those with low WMC regardless of the memory load (Hypothesis 2B). Most importantly, levels-of-processing and memory load had a marginally significant interaction (Hypothesis 2C). While shallow and moderate processing were subject to the effects of memory loads (i.e., fewer words were recalled under memory load conditions), deep processing appears to be robust enough to overcome the effects of memory taxation. The implications of these findings are somewhat diminished since the predicted interaction between WMC groups and processing were not found; however, this result does suggest that encoding strategies that incorporate deep processing remain successful even under periods of increased competition for cognitive resources. Whether this phenomenon persists under other conditions that inhibit cognitive processing (e.g., stress) may be worth investigating in future studies.

Although WMC group was not found to interact with processing, a statistically significant three-way interaction was found between WMC group, processing, and memory load. Follow-up analyses, however, reveal a perplexing pattern of results (See Figure 3 and Figure 4). Both the high and low WMC groups
showed gradual decline in the number of words recalled as memory load increased at both shallow and moderate processing. The low WMC group had statistically significant differences at deep processing, while the high WMC group performed significantly poorer with no memory load while processing words at the deepest level. Theoretically speaking, the high WMC group should have performed at its apex when utilizing the most advantageous processing strategy in the absence of a memory load. Therefore, the only safe conclusion is to assume that subjects with high WMC underperformed in this condition of the experiment. Thus, the three-way interaction may be the result of a type-one error.

Secondary analyses strengthen this study, as extraneous variables (e.g., age, ethnicity, time of day, experimenter) were not found to alter the number of words reported by subjects. The same was true for operation-span scores in relation to most of the previously mentioned variables with the exception of the experimenter variable. Further complicating this finding is the fact that each experimenter worked on specific days of the week throughout the period of data collection. The experimenter with the highest mean worked exclusively on Tuesdays and Thursdays, while the second experimenter worked Tuesdays and Fridays, and the experimenter with the lowest mean worked exclusively on Mondays and Wednesdays. There were no outliers among the participants that explain the differences between experimenters, as the experimenter with the highest mean simply had a higher frequency of subjects score above 50. It should be noted that the only day that was unique to this experimenter was Thursday, which also serves as the end of the typical college
student’s week. Improved mood may have resulted in overachievement in some subjects.

All experimenters worked from a loose script that kept interaction with participants to a minimum, as all detailed instructions for the two experiments were written into the E-prime file used to present the stimuli. All experimenters had hands-on experience with the E-prime experiments for the purpose of responding to questions subjects may have had in the process of completing the tasks. If differences between the experimenters did have an effect on operation-span performance, they may have arisen in their approach to fielding questions. It is also the case that the experimenter with the highest average was nearest to the mean age of the participants. This may have increased her likeability, resulting in increased motivation to perform on the task. It is equally perplexing to decipher how the day of the week may have interfered with task performance. There is no obvious reason to expect that persons with high WMC were most likely to participate in this study on Tuesdays and Thursdays.

Both elements of this study (i.e., the levels-of-processing task and the operation-span) were lengthy, demanding cognitive tasks. It was important to test whether or not performance dwindled as time spent on task (i.e., the number of blocks the participant had seen) increased. Based on the result of this analysis, it can be assumed that fatigue was not a factor. It was also of interest to test how increased exposure to different stimuli sets factored into interference (i.e., words recalled from stimuli sets seen prior to the focal block). Previous studies suggest that individuals
with high WMC should experience less interference (Unsworth, 2007). Based on the results, it can be concluded that increased exposure to sets of stimuli does increase interference. There was, however, no significant difference between high and low WMC groups in terms of the number of words reported from previous sets. This failure to replicate is somewhat troubling, as it represents yet another instance when high WMC individuals failed to outperform their low WMC counterparts.

Due to the frequency with which analyses that have assessed the WMC groups have failed to find significance, it may be worth speculating about how effectively this study distinguished high from low WMC individuals. Most concerning is the observed effect of the experimenter on operation-span scores. Whether this effect is explained by the likability of a specific experimenter or by the day of the week subjects participated in this study, it may be the case that there were individuals whose performances would more often reflect average (or slightly above average) WMC grouped in with high WMC individuals. The next question that must be asked is whether or not inclusion of such individuals would weigh down group averages such that any differences between high and low WMC would no longer detectable. Use of the extreme quartile design is undertaken to defend against this particular problem. Even if subjects overachieved, it is unlikely that someone in the below average group could manage their way into the high WMC group. It may also be worth repeating that the descriptive statistics of operation-span scores in this study (independent of experimenter) mirror those of samples collected prior to the current study (IlIingworth & Schustack, 2010; Unsworth, Heitz, Schrock, & Engle, 2005). It
is therefore difficult to conclude that this study was flawed due to the way in which the grouping variable was formulated.

Although no statistical analysis could be calculated, data concerning performance in response to each of the six sets of stimuli were reported (Table 3). It should be noted that the randomization of the stimuli sets to condition worked reasonably well as the number of subjects to see each set were nearly equal, ranging from 54 to 66. Similarities between the stimuli sets in terms of means, medians, standard deviations, and ranges suggest that stimuli sets did not factor into the experimental outcomes.

**Conclusion**

Overall, this study fell short of shedding light on the nature of WMC and the underlying cognitive processes that may account for differences found across the population. This ranged from a failure to detect any differences in rates of free recall, responses to different levels-of-processing, responses to increased memory load, and interference. Despite the limitations of the current study, the size of the effect typically found between samples on extreme ends of the WMC spectrum suggests that the levels-of-processing paradigm may not be suitable for investigating differences in processing or encoding that occur in these populations, if they do, in fact, occur. Future studies may wish to incorporate new control mechanisms for processing information.

This study did, however, demonstrate that the effect of memory loads on free recall is different across the various levels-of-processing. Given the complex way in
which WMC interacted with this relationship, it may be the case that this phenomenon is applicable to everyone, regardless of any naturally occurring cognitive limitations. It may be of interest to explore the generalizability of this effect, as there are a variety of real-world situations where environmental factors may act on human cognition as a memory load. A related avenue of research would be to examine how likely (or able) individuals are to utilize such processing strategies under such environmental stress.
References


a probability judgment paradigm. Poster session presented at the annual meeting of the Society for Personality and Social Psychology, San Antonio, Texas.


Lespinet-Najib, V., N’Kaoua, B., Sauzeon, H., Bresson, C., Rougier, A., & Claverie


Neuroscience, 3, 85-90.


Figure 1. Order of Events. This figure illustrates the series of slides seen by subjects in each block.
Figure 2. Words recalled over increased memory load. This figure illustrates the nature of the interaction between memory load and levels-of-processing.
Figure 3. Low working memory capacity group. This figure illustrates how levels-of-processing performance changed over increasing memory load for individuals who scores in the lowest quartile of the operation-span task.
Figure 4. High working memory capacity group. This figure illustrates how levels-of-processing performance changed over increasing memory load for individuals who scores in the highest quartile of the operation-span task.
Figure 5. Interference over time. This figure illustrates the linear increase in interference as the number of blocks seen by subjects increases.
Figure 6. Words recalled over increased level of processing. This figure illustrates the rate at which words recalled increases, as processing deepens for both the high and low WMC groups.
Table 1

Means and Standard Deviations for O-span Samples

<table>
<thead>
<tr>
<th>Current</th>
<th>Illingworth &amp; Schustack, 2010</th>
<th>Unsworth et al., 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>120</td>
<td>36.38</td>
<td>17.22</td>
</tr>
</tbody>
</table>
Table 2

WMC Group O-span Statistics

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>14.53</td>
<td>4.95</td>
<td>3</td>
<td>22*</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>29.27</td>
<td>3.61</td>
<td>22*</td>
<td>34</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>42.47</td>
<td>3.54</td>
<td>36</td>
<td>48</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>59.27</td>
<td>6.30</td>
<td>50</td>
<td>70</td>
</tr>
</tbody>
</table>

Note. Subjects scoring 22 on the o-span were split between groups 1 and 2 based on the total number of letters they reported in the correct spot.
Table 3

Descriptive Statistics for Stimuli Sets

<table>
<thead>
<tr>
<th>Stimuli set</th>
<th>N</th>
<th>Mean</th>
<th>Median*</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1</td>
<td>55</td>
<td>12</td>
<td>12</td>
<td>5.03</td>
<td>2</td>
<td>26</td>
</tr>
<tr>
<td>Set 2</td>
<td>62</td>
<td>9.92</td>
<td>10</td>
<td>5.68</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>Set 3</td>
<td>66</td>
<td>9.97</td>
<td>9.50</td>
<td>4.32</td>
<td>2</td>
<td>26</td>
</tr>
<tr>
<td>Set 4</td>
<td>60</td>
<td>9.75</td>
<td>9.50</td>
<td>4.42</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>Set 5</td>
<td>54</td>
<td>10.82</td>
<td>10.50</td>
<td>5.09</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>Set 6</td>
<td>62</td>
<td>10.82</td>
<td>10.50</td>
<td>5.45</td>
<td>1</td>
<td>33</td>
</tr>
</tbody>
</table>

*Note. Medians were reported due to the skewed distribution of these variables.*