12-2011

Change Detection Memory In Rhesus Monkeys And Humans

Lauren C. Elmore

Change Detection Memory in Rhesus Monkeys and Humans

By

Lauren Caitlin Elmore, B.S.

APPROVED:

_____
Anthony A. Wright, Ph.D.
Supervisory Professor

_____
Michael Beauchamp, Ph.D.

_____
Timothy Ellmore, Ph.D.

_____
Richard Meisch, M.D., Ph.D.

_____
Anne Sereno, Ph.D.

APPROVED:

_____

Dean, The University of Texas
Health Science Center at Houston
Graduate School of Biomedical Sciences

CHANGE DETECTION MEMORY IN RHESUS MONKEYS AND HUMANS

A

DISSERTATION

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
and
The University of Texas
M.D. Anderson Cancer Center
Graduate School of Biomedical Sciences
In Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Lauren Caitlin Elmore, B.S.
Houston, TX

December, 2011

DEDICATION

This work is dedicated to Cisco and Captain.

ACKNOWLEDGEMENTS

I would first like to acknowledge my advisor, Dr. Anthony Wright for his guidance over the past five and a half years. I would also like to thank him for believing in me when I was an an undergraduate summer student and encouraging me to pursue this Ph.D. I cannot imagine myself in any other career and truly have him to thank for his encouragement. I am also grateful that he pushed me to work hard and think critically about science while simultaneously allowing me the freedom to pursue research that I am passionate about.

Secondly, I would like to thank my committee members past and present for their guidance and support. I am particularly grateful to my supervisory committee members; Dr. Mike Beauchamp for teaching me to think on my feet, Dr. Anne Sereno for emphasizing the importance of clarity and organization in research, and to Drs. Tim Ellmore and Richard Meisch for the friendly encouragement. I would also like to acknowledge Drs. Pramod Dash, Valentin Dragoi, Andrew Papanicolau, and Jack Waymire for their guidance in the early stages of my graduate career.

I feel considerable gratitude for the guidance of Jacquelyne Rivera, Research Associate, throughout much of my graduate career. She taught me nearly everything I know about working with rhesus monkeys, and her love for these animals was infectious. She showed me how to treat the animals in order to have healthy and happy research subjects. Her hard working nature and industriousness is second to none and it would have been difficult to accomplish many of the experiments discussed here without her invaluable assistance.

I also owe many thanks to our collaborators. I am extremely grateful to Dr. Wei Ji Ma for his guidance with modeling visual short term memory. Without his help, the theoretical aspects of this dissertation would have been extremely difficult to accomplish. I am thankful to John Magnotti for his assistance with statistics and analysis, in particular his help with the

multidimensional scaling analysis. I am also for grateful to Dr. Jeffrey Katz for his incredible statistical knowledge, creative research ideas, and continual encouragement over the years.

I would also like to thank my labmates past and present for making the laboratory an enjoyable work environment. I thank Tony Passaro for his help with the human studies presented here and many good discussions of the visual short term memory literature. I thank Dr. Kenneth Leising for being such a pleasant and fun person and a good role model for what a scientist should be. Lastly, I thank Deepna Devkar for her help and encouragement especially during the time period in which I wrote this dissertation.

I would next like to thank the CLAMC staff for their care of and assistance with the animals over the year. In particular I would like to thank Drs. Bradford Goodwin, Vanessa Jensen, and Chris Smith as well as Jackie Bludworth, Carol Stavinoha, Jimmy Soto, Devielle Bell, and Derrick Hagler for helping keep the animals healthy and happy.

Finally, I am very appreciative for the support that my friends and family have provided over the years. My parents in particular have been amazing to me throughout my graduate career. Without the support of the people close to me, this would have been a very difficult journey.

ABSTRACT

CHANGE DETECTION MEMORY IN RHESUS MONKEYS AND HUMANS

Publication No. _____

Lauren Caitlin Elmore, B.S.

Supervisory Professor: Anthony A. Wright, Ph.D.

Visual short-term memory (VSTM) is the storage of visual information over a brief time period (usually a few seconds or less). Over the past decade, the most popular task for studying VSTM in humans has been the change detection task. In this task, subjects must remember several visual items per trial in order to identify a change following a brief delay interval. Results from change detection tasks have shown that VSTM is limited; humans are only able to accurately hold a few visual items in mind over a brief delay. However, there has been much debate in regard to the structure or cause of these limitations. The two most popular conceptualizations of VSTM limitations in recent years have been the fixed-capacity model and the continuous-resource model. The fixed-capacity model proposes a discrete limit on the total number of visual items that can be stored in VSTM. The continuous-resource model proposes a continuous-resource that can be allocated among many visual items in VSTM, with noise in item memory increasing as the number of items to be remembered increases.

While VSTM is far from being completely understood in humans, even less is known about VSTM in non-human animals, including the rhesus monkey (*Macaca mulatta*). Given that rhesus monkeys are the premier medical model for humans, it is important to understand their VSTM if they are to contribute to understanding human memory. The primary goals of

this study were to train and test rhesus monkeys and humans in change detection in order to directly compare VSTM between the two species and explore the possibility that direct species comparison might shed light on the fixed-capacity vs. continuous-resource models of VSTM. The comparative results suggest qualitatively similar VSTM for the two species through converging evidence supporting the continuous-resource model and thereby establish rhesus monkeys as a good system for exploring neurophysiological correlates of VSTM.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER 1: GENERAL INTRODUCTION

Visual short-term memory (VSTM) is the mental storage of visually presented information over a brief delay of a few seconds or less. VSTM is undoubtedly a major mechanism by which humans (and animals) maintain awareness of their constantly changing environments, and behave adaptively. For humans, VSTM is important for tasks like driving, wherein we must act based on both our current views of the road conditions and other cars, but also based on our memory for what was seen in the sideview mirror a few seconds ago, for example. Likewise, non-human animals also use VSTM to detect changes in their environment. Maintaining constant vigilance is impossible, as animals must engage in other behaviors, such as foraging, caring for young, and building habitats. Consider the example of a rhesus monkey, searching for fruit in the tree tops of an Indian forest. The monkey must focus his attention on finding fruit but must also be vigilant for potential predators, such as raptors. If the monkey checks his surroundings at regular intervals, it may notice a change in the shadow on a nearby tree. Such a change could help alert the monkey to the presence of a raptor perched atop the tree, allowing time for escape. Noticing the shadow would depend on memory of the tree's prior appearance.

Aside from its role in the safety and survival of humans and other animals, short-term memory is an important component in memory generally, given its role as the gateway to long-term memory (Modal model of Atkinson & Shiffrin, 1968). Incoming visual information is at first transiently stored in VSTM before being consolidated into long-term memory (Fukuda & Vogel, 2010; 2011). Again, the driving and fruit searching examples highlight the importance of visual long-term memory. For the driver, it is important for him/her to remember that the fork in the road marked by a pecan tree is the place where he must veer right in order to complete the journey to his summer cabin. This information about the pecan tree entered his memory via VSTM the first time he made the trip, and was ultimately consolidated into long-term memory, allowing him/her to remember this tree year after year. The monkey must use visual cues to remember where he found fruit. After he and his fellow monkeys deplete one tree of fruit, it is

advantageous for them to remember that tree's location and appearance over the long term as they will want to revisit once more fruits become ripe. Again information about the tree's appearance must pass through VSTM before it is consolidated and more permanently stored in long-term memory.

A scientific investigation of the mechanisms of VSTM is important for many reasons. Considering the important role of VSTM in daily life and its connection to long-term memory, it is often quite debilitating when VSTM fails. Interestingly, short-term memory is impaired in numerous psychological and neurological disorders, including: Alzheimer's disease, Parkinson's disease, Huntington's disease, Tourette's syndrome, Dementia with Lewy bodies, Progressive Supranuclear Palsy, Traumatic Brain Injury, Stroke, Multiple Sclerosis, Attention Deficit-Hyperactivity Disorder, Schizophrenia, Depression, & Post-Traumatic Stress Disorder (e.g, Baddeley et al., 1986; Baddeley et al., 1991; Dubois & Pillon, 1997; Gabrieli, 1998; Budson & Price, 2005; Brandes et al., 2002, Koenen, et al., 2001). It is therefore important to understand VSTM when it is functioning normally in order to understand how various disease states impair VSTM.

Over the course of the past two decades, the change detection task has become an increasingly popular procedure for the study of VSTM (e.g., Alvarez & Cavanagh, 2004; Eng, Chen, & Jiang, 2005; Luck & Vogel, 1997; Pashler, 1988; Rensink, 2002; Wilken & Ma, 2004). In this task, a sample display of two or more stimuli is presented. Following a brief delay (e.g., one second) the subject has to judge which item changed or whether or not a change occurred. Studies using change detection have shown that visual short-term memory performance declines as a function of display size (the number of stimuli presented in the sample display (e.g., Alvarez & Cavanagh, 2004; Eng, Chen, & Jiang, 2005; Luck & Vogel, 1997; Pashler, 1988; Wilken & Ma, 2004). This is an expected result; the task becomes more difficult as the display size increases because there are more items to remember. This result does however highlight an important fact, that VSTM is limited. The system is not capable of

processing and storing unlimited numbers of visual items. This begs the question of how VSTM works when it is overwhelmed with information to store? Interestingly, during the past decade two competing models were developed that seek to explain the limitations of VSTM. These models, the fixed-capacity model (Cowan, 2000; 2005) and the continuous-resource model (Wilken & Ma, 2004; Bays & Husain, 2008) suggest very different mechanisms for VSTM function.

The fixed-capacity model has its roots in George Miller's seminal 1956 paper, "The magical number seven, plus or minus two." Although this paper was less about precise memory storage limits and more about the ability to increase storage capacity through purposeful grouping of items (chunking), it was the first to suggest that short-term memory has a limit which can be quantified. Since the publication of Miller's paper, other possibilities have been proposed to explain the limitations of VSTM. Some authors argued that the limitation is in the duration of time in which an item can be stored in short-term memory without being actively rehearsed (e.g. Baddeley, 1986; Sperling, 1963). Cowan previously suggested that the storage mechanism is time limited whereas the focus of attention is capacity-limited (Cowan, 1988; 1995). However, in the past decade, the overwhelmingly popular model mechanism of VSTM has been Cowan's more recent "fixed-capacity model" which states that VSTM is a capacity-limited storage mechanism which on average (across subjects) consists of "slots" for three to five items (Cowan, 2000; 2005). Cowan notes that this limit, or "magic number 4±1" is applicable to short-term memory in multiple modalities, including visual, verbal, and auditory short-term memory. Cowan reviews converging evidence supporting the "magic number 4±1" come from various tasks including visual partial vs. whole report (e.g. Sperling, 1960), auditory whole report (e.g., Darwin et al., 1972), whole report of spoken lists (e.g. Baddeley, 1986), in addition to change detection. It should be noted that Cowan was not the first or only researcher to propose a fixed capacity of VSTM (e.g., Alvarez & Cavanagh, 2004;

Eng, Chen, & Jiang, 2005; Luck & Vogel, 1997; Pashler, 1988; Vogel et al., 2001), but has been the most prolific in his description of the model.

There are essentially two schools of thought within this group of memory researchers investigating the fixed-capacity model. One group has proposed a more rigid interpretation of the model, in which the "magic number 4±1" is constant across stimulus types and visual complexity. For instance, Luck & Vogel (1997) found that the capacity limit of 4±1 was constant despite increases in the information that had to be maintained for each stimulus. In this experiment subjects were instructed to remember colored bars of varying orientation. In one condition subjects were instructed to only remember the colors, in another condition only the orientations, and finally in the most complex condition to remember both the color and the orientation. They found that the capacity limit did not decrease when subjects had to remember both color and orientation, a finding indicating that complexity of the visual information stored does not influence VSTM capacity.

The second school of thought takes a more flexible interpretation of the fixed-capacity model and has shown that capacity could be a function of stimulus complexity (e.g., Alvarez & Cavanagh, 2004; Eng, Chen, & Jiang, 2005). Alvarez & Cavanagh (2004) and Eng et al. (2005) tested subjects in both change detection and visual search tasks using a variety of stimulus types including colored squares, shaded cubes, squiggles, faces, random polygons, Snodgrass line drawings, and Chinese characters. They found that stimulus complexity (as determined by visual search time) had an inverse correlation with VSTM capacity for that stimulus type. For instance, Eng et al. (2005) showed that capacity was highest with alphabet letters and lowest for human faces. Eng et al. (2005) found a weaker correlation between visual search time and capacity (r^2 = 0.76) than did Alvarez & Cavanagh (2004) as they reported an r^2 value of 0.992. It is odd to claim that a discrete-slot storage system should vary in capacity (change the number of memory slots) based on stimulus type, Eng et al. (2005) showed that increasing the viewing time reduced the variance in capacity measures across

stimulus types. They proposed that the differences in capacity limits across stimulus types were the result of a perceptual bottleneck in which more complex objects required longer encoding time in order to be adequately stored in VSTM. Alvarez & Cavanagh interpreted their similar results in a more complex conceptualization of VSTM: they hypothesized that there is a maximum visual information limit, which is the product of the number of stimuli and the visual information per stimulus, and that there is a secondary limit in that the maximum number of stimuli maintained cannot be greater than four or five.

Cowan's (2000) description of the fixed-capacity model sparked several attempts to investigate the neural correlates of the "magic number 4±1". For instance, Todd & Marois (2004) demonstrated using functional magnetic resonance imaging (fMRI) that brain activity in bilateral intraparietal sulcus increased as the memory load (number of items to remember) increased. They claimed that this finding suggests that the locus of VSTM lies in the posterior parietal cortex. They also found that brain activity was maximal during the delay period when subjects had three or four items to remember and that this brain activity did not increase further when the memory load was as large as eight items. The authors interpreted these results as evidence that the intraparietal sulcus tracks VSTM capacity, and proof positive for the notion of a capacity limit of 4±1.

In a study in the same issue of *Nature,* Vogel & Machizawa (2004) claimed that VSTM capacity varied across a large range among subjects (from 1.5 to 5 objects), but that they too had identified a neural correlate of a fixed-capacity, in this case using event-related potentials (ERPs). They found that the contralateral ERP signal during the retention delay (presumed to originate from posterior parietal and lateral occipital regions) was maximal at the VSTM capacity of the individual subject. For instance, an individual with a low capacity of two would have maximum ERP signal for memory loads of two or greater. Likewise, an individual with a larger capacity of four items would have maximum ERP signal for memory loads of four or greater. Thus, the ERP signal was tied to the individual subject's capacity limit. However,

what remains unclear is whether or not a maximum neural signal at a particular capacity value is definitive evidence for a fixed-capacity limit. Neither of these papers provided much information on within subject variability across trials and conditions.

A third investigation of the neural substrates of VSTM capacity was pursued by Buschman et al. (2011). They trained rhesus monkeys in a change detection task and recorded from neurons in three brain regions, lateral intraparietal cortex, lateral prefrontal cortex, and the frontal eye fields. Based on their behavioral and electrophysiological results, Buschman et al. (2011) concluded that there is a capacity limit, and also that the right and left hemifields have independent capacity limits. These authors suggested that their electrophysiological data indicated that the information bottleneck giving rise to capacity limits, originates in posterior parietal cortex. However, these authors noted that increasing the memory load (more items to remember) reduced the information available about each individual stimulus in all three brain regions. It is unclear how this finding could support a slot-like fixed-capacity memory because slot-like representations must be all or none. The findings of Buschman et al. (2011) seem to suggest a continuous allocation of memory resources.

The continuous-resource model of Wilken & Ma (2004) is an approach to VSTM that is more closely tied to what is known about computations in the brain in general. Wilken & Ma (2004) began their description by pointing out the unattractive features of fixed-capacity models. Of particular interest was that fixed-capacity models proclaim that individual stimuli are encoded all-or-nothing within the brain. Specifically fixed-capacity models claim that an item is either present or absent, and there is no noise in the system. Wilken & Ma (2004) argued that discrete noise-free representations are neurally implausible. Instead, they proposed VSTM to be a continuous variable (rather than discrete units) which can be modeled using signal detection theory. In place of computing the discrete variable of capacity, they used *d'* from signal detection theory to quantify memory sensitivity. The model predicts that like accuracy, *d'* should fall as the memory load increases, but also that this decline in *d'*

should be fit by a power law function, because *d'* should fall as a function of 1/n, where n is the number of items in the memory display. Although memory as a continuous-resource can be allocated to large numbers of stimuli, performance and *d'* fall because the noise in the memory representation also increases with larger displays of stimuli. Wilken & Ma (2004) also noted that fixed-capacity models are overly complex and that the simple assumption of neural noise along with the simple decision rule from signal detection theory is sufficient to understand VSTM. They also suggested that the fixed-capacity model "magic number 4±1" has received considerable empirical support due to an artifact of increasing noise as the memory load increases.

Modeling VSTM as a continuous-resource has received support from some recent studies (Bays et al., 2009; Bays & Husain, 2008). In 2008, Bays & Husain showed that memory is a resource which is shared between all objects in a scene, but can be flexibly shifted and weighted more heavily toward certain stimuli based on selective attention. They argued that the precision of storage of an item depends on the total number of items to be stored. In a second article, Bays et al. (2009) showed that the continuous-resource model could account for findings that had been previously argued to support a modified fixed-capacity slot model (Zhang & Luck, 2008).

Thus, it seems that both the fixed-capacity model and the continuous-resource model both have received support in the literature based on results from human subjects. One approach to disambiguating hypotheses in human cognition is to see if the same hypotheses can garner support in non-human animals. Because cognitive processes such as VSTM are subject to evolution the same or similar cognitive processes may be present in our ancestral species. This is more likely to be the case is closely related species, such as non-human primates, but can also occur in more distantly related species.

The approach of comparing across species both closely related (humans and rhesus monkeys) and more distantly related (pigeons) was taken with great success by Wright and

colleagues (1985). They were interested in understanding the mechanism of U-shaped serial position functions in visual list memory. Serial position functions had been known to occur in human memory, and the primacy and recency effects shown in these functions were often attributed to verbal rehearsal processes (e.g., Atkinson & Shiffrin, 1968). However, in this seminal *Science* paper, Wright and colleagues (1985), showed that serial position functions were present in humans, rhesus monkeys, and pigeons, and that they primarily differed in time course, with the changes from recency to primacy occurring fastest for pigeons and slowest for humans. The finding of primacy and recency effects in monkeys and pigeons eliminated the possibility that these effects were due to verbal rehearsal, and demonstrated that the underlying mechanism must qualitatively be the same across species. Serial position functions have been shown in other species including capuchin monkeys (Wright, 2007), apes (Buchanan et al., 1981), squirrel monkeys (Roberts & Kraemer, 1981), and rats (Bolhuis & van Kampen, 1988; Harper, McLean, & Dalyrmple-Alford, 1993; Kesner & Novak, 1982; Reed et al., 1996), thereby providing converging evidence that serial position effects occur in recognition memory in a variety of species, and thus that the cognitive process underlying recognition memory has been conserved throughout much of recent evolution.

Given the example from serial position functions, it is easy to imagine how comparing the two models of VSTM across species may provide converging evidence in support of either the fixed-capacity or the continuous-resource model. Testing animals with the change detection task and finding that they too have a fixed capacity or "magic number" which could also be four but might differ (e.g., smaller) would provide further evidence in favor of the fixed-capacity model. Likewise, finding that memory sensitivity (*d'*) can be characterized by power law functions in multiple species would provide support for the continuous-resource model.

Testing non-human animals in change detection has other advantages as well. Not only is understanding animal memory interesting in its own right, but also, animals provide opportunities to conduct studies that cannot be done in humans for practical reasons. Animals

can be tested for many more sessions than would be practical with human subjects. Large numbers of sessions increase the statistical power of the study and provide a more stable measure of the construct of interest, in this case VSTM. Greater control is possible with animal studies, as the subjects can easily be tested at the same time daily with greater control over the subject's motivational state. Since animals routinely work for food or liquid reinforcement, and this can be tightly controlled by the experimenter. Furthermore, non-human animals ultimately allow the direct manipulation and investigation of the neural substrates of cognitive processes through invasive studies including lesions, electrophysiological recordings, inactivation, stimulation, and pharmacological and neurotransmitter manipulations (although it should be noted that this type of work is beyond the scope of this dissertation).

An obvious first choice of non-human species to compare with human subjects in the change detection task is the rhesus monkey (*Macaca mulatta*). Rhesus monkeys are the premier medical model for humans due to their highly similar genetics, anatomy, and physiology (Rhesus Macaque Genome Sequencing and Analysis Consortium, 2007). They are closely related to humans and diverged from a common ancestor in the relatively short evolutionary timescale of 25 million years ago (Kumar & Hedges, 1998). Their neuroanatomy is well known and is similar to humans.

Like humans, monkeys are particularly predisposed to visual tasks like change detection because they have a highly developed visual system. Both species have a large percentage of their cortex devoted to vision: 50% in rhesus monkeys and 30% in humans (Van Essen, 2004). Another critical brain region for VSTM is the prefrontal cortex, as demonstrated by electophysiological, lesion, and inactivation studies in rhesus monkeys (e.g., Funahashi et al., 1989; Fuster & Bauer, 1974; Petrides, 1994; 1996; Sawaguchi & Goldman-Rakic, 1991; Wilson et al., 1993) and human neuroimaging studies (e.g., Pessoa & Ungerleider, 2004; Sala & Courtney, 2007). A direct comparison of the architecture of the prefrontal cortex between

monkeys and humans has shown that the architectonic organization is quite similar (Petrides, 2005). In both species, the cortex is organized along a rostral-caudal axis as well as a dorsal ventral axis. Functional roles for the different subregions of the prefrontal cortex appear to correspond between monkeys and humans, including the regions involved in motor control, cognitive control, working memory, and decision making. Given the qualitative similarities and close anatomical organization of both monkey and human visual and prefrontal cortices, one might predict that monkey and human performance in a VSTM task should be qualitatively similar.

However, despite the striking similarities between humans and monkeys there are quantifiable differences in the neuroanatomical substrates of VSTM that would likely give rise to quantitative differences in absolute performance levels. This hypothesis is in accord with the Darwinian perspective, that cognitive differences between animals and humans are a matter of degree and not of kind (Darwin, 1872). For one, rhesus monkey brains are both smaller than human brains: and more specifically their brain to body mass ratio or encephalization quotient is smaller. For humans the average encephalization quotient is 7.4-7.8 and for monkeys it is 2.1(Roth & Dicke, 2005). However, differences in encephalization quotient should be interpreted with caution, other studies have suggested that overall brain size is more important (e.g. Deaner, et al., 2007), and it is unclear how much variance in cognitive function can be related to encephalization quotient or total brain size.

The human prefrontal cortex though architecturally similar to that of the rhesus monkey is much larger and occupies a greater proportion of the brain (Semendeferi et al., 2002). Another quantifiable difference relates to the pyramidal cells of the prefrontal cortex. By comparing pyramidal cell morphology across humans, rhesus monkeys, and marmosets, Elston et al. (2001) demonstrated that pyramidal cells in the prefrontal cortex have become more branched and spined over the course of evolution (human cells were the most branched and spined), allowing greater numbers of connections between neurons. Elston et al., (2001)

suggested that this difference in neuronal morphology likely underlies the more advanced cognitive capabilities that occur in humans. In general, the differences in prefrontal cortex size, neuron morphology, and neuronal density are thought to underlie the differences in cognitive abilities presumed to exist between humans and other mammals (Roth & Dicke, 2005). Thus, it seems that rhesus monkeys have a similar enough cortical architecture that their VSTM may be qualitatively similar to humans, but that there are enough differences, particularly in the prefrontal cortex, that one would expect a quantitative difference between the two species.

There were four main goals in this study. We began our research (Chapter 2) by testing human subjects in change detection in order to determine if our paradigm would yield results similar to those previously published. We also tested human subjects with the parameters which would eventually be used with rhesus monkeys in order to allow a direct comparison. In a second experiment, we investigated the role of stimulus type in VSTM by testing various types of stimuli.

The second goal was to train rhesus monkeys to perform the change detection task (Chapter 3). Because this is one of the most popular tasks for the study of VSTM in humans, it would be advantageous to determine if the task could be learned by monkeys. If monkeys could learn change detection, would the monkeys perform change detection in an analogous way to humans? Such a result would validate comparisons between the two species.

The third goal was to compare VSTM between humans and monkeys using the change detection task. Most animal memory procedures are simplistic (e.g. delayed match to sample) relative to recall and recognition procedures used with humans. As such, training rhesus monkeys in change detection provided a unique opportunity to directly compare memory (and its limitations) across species using identical or nearly identical task parameters. In Chapter 4, we tested rhesus monkeys in a task similar to that used to test humans (Chapter 2) in order to determine if VSTM was qualitatively similar between species. In Chapter 5, we extended the

study began in Chapter 4, by testing rhesus monkeys with the same stimuli, viewing time, delay period, and many of the same display sizes used with humans in Chapter 2. To provide a more direct comparison of VSTM between rhesus monkeys and humans, in Chapter 5 the test parameters were made as similar as possible given the constraints of each species.

Finally, the fourth goal of this dissertation was to use the results generated in Chapters 2, 4, and 5 to guide a more thorough theoretical understanding of VSTM function, and determine whether or not these results and comparisons could discriminate between fixed-capacity and continuous-resource models of VSTM. Specifically, we compared the fixed-capacity and continuous-resource models to determine which model provided a better fit to the data, and which model was more theoretically sound given the results that we found.

CHAPTER 2: VISUAL SHORT-TERM MEMORY IN HUMANS

Introduction

Despite decades of research, VSTM is only beginning to be understood. The field is riddled with conflicting interpretations, only coming together in agreement that VSTM is limited because memory accuracy declines as a function of the display size, or the number of stimuli that one must remember (e.g., Alvarez & Cavanagh, 2004; Eng, Chen, & Jiang, 2005; Luck & Vogel, 1997; Pashler, 1988; & Wilken & Ma, 2004). In fact, the limitations in short-term memory have been demonstrated for visual, verbal, and auditory information. VSTM storage limitations are in contrast to long-term visual memory, where research has shown extremely large storage capacities for pictures, words, and associations (Standing, et al., 1970; Standing, 1973; Shepard, 1967; Voss, 2009).

As mentioned in the General Introduction, the greatest point of contention in the VSTM arises from the dispute between competing models, the fixed-capacity model and the continuous-resource model. Both models seek to provide structure and a functional basis to the inherent limitations in VSTM. Differences between these two approaches are further developed here. The work of Luck & Vogel (1997) and Cowan (2001) has supported modeling VSTM as a slot-like storage system of a limited (and fixed) capacity. The model has a discrete number of slots for the storage of visual information. The popularized number of slots is the so-called "magic number 4±1." A computational model has also been developed to explain this magic number (Rouder et al., 2008). Figure 2.1 provides a visual conceptualization of how fixed-capacity models are thought to work.

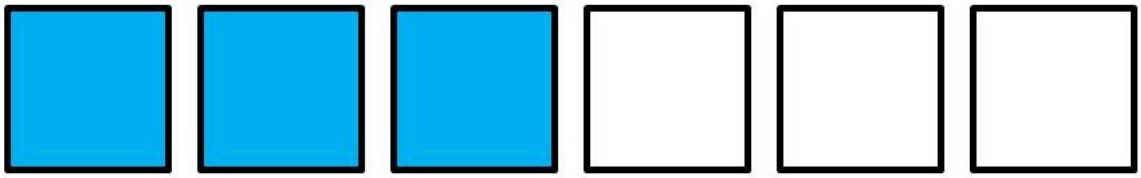


Figure 2.1: Fixed-Capacity Memory System with three slots. Black box outlines represent visual stimuli to be remembered. If an individual has a VSTM capacity of three, then six exceeds the capacity limit by three items. Blue fills represent memory. According to a fixed-capacity model memory (blue) is allocated according to the number of slots available (three). As a result three stimuli are perfectly stored, and the other three are subsequently forgotten.

The apparent simplicity of a fixed number of memory slots has been muddled by some researchers supporting this view. For instance, Brady et al. (2011) and Alvarez & Cavanagh (2004) suggested that VSTM should be characterized both by the number of items it can store (capacity) but also by the fidelity of storage. This emphasis on fidelity has arisen from findings demonstrating that VSTM capacity differs based on stimulus type (e.g. Alvarez & Cavanagh, 2004; Eng et al., 2005). While one group (Eng et al., 2005) indicates that stimulus differences in capacity can be eliminated with additional viewing (study?) time, Alvarez & Cavanagh (2004) proposed that to adequately model VSTM, one must create a model that incorporates both visual information (stimulus complexity) and a limited number of storage slots.

Rouder et al. (2008) added yet additional parameters of attention to "salvage" the fixed-capacity model. In lieu of adding parameters to rescue the fixed-capacity model,is it possible to account for the findings of VSTM experiments with a simpler solution?

A promising and simpler model of VSTM limitations is the continuous-resource model of Wilken & Ma (2004). Rather than modeling memory as a discrete entity of a few slots, the continuous-resource model states that memory is a continuous-resource that can be allocated to *many* stimuli. A visual conceptualization of this is displayed in Figure 2.2.

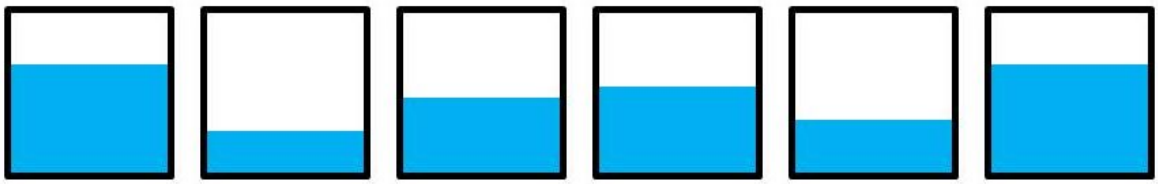


Figure 2.2: Continuous-Resource Model of VSTM. Black box outlines represent visual stimuli to be remembered. Blue fills represent memory. Memory is distributed among all items, but there is not enough resource to perfectly store all six stimuli.

Instead of proposing a discrete limit in terms of number of stimuli (capacity), the continuous-resource model proposes that memory should be distributed among all stimuli, but with increasing numbers of stimuli there is less resource per stimulus. A reduction in resource per stimulus with increasing display size results in increasing noise in those memory representations. The fidelity of the representation of each item is decreased. An obvious extension is that complex stimuli would require more resource per stimulus, thereby resulting in the performance differences based on stimulus type, a result found in previously discussed studies (e.g. Eng et al., 2005; Alvarez & Cavanagh, 2004).

The continuous-resource model is based on signal-detection theory (e.g. Green & Swets, 1996, & Macmillan & Creelman, 2005), the predominant theory for how discriminations are made. The model sees VSTM as a matter of discriminating memory for an item from noise. The model uses *d'* from signal detection theory as a measure of memory sensitivity (Figure 2.3). In change detection, the model posits that each stimulus in the sample display is represented with noise in memory. The noise in the memory representation can lead to perceived changes in both test stimuli (although only one has in fact changed). In order to make a decision as to which stimulus changed, the subject must compare the perceived changes to zero. The probability of each item being the changed stimulus is represented by a Gaussian curve and *d'* or memory sensitivity corresponds to the ability to distinguish between the two curves and identify the item which has actually changed.

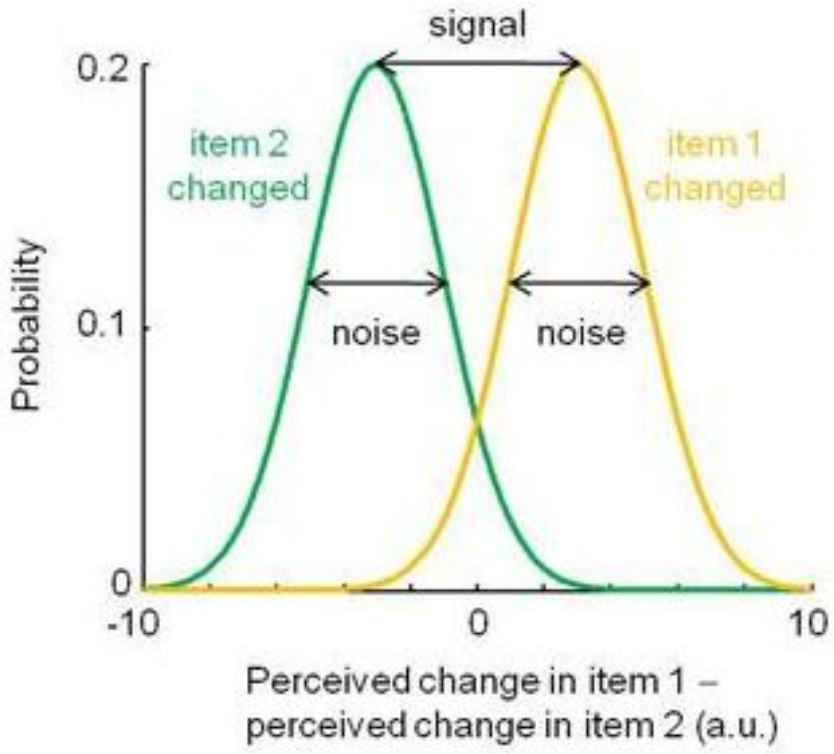


Figure 2.3: Memory Sensitivity – *d'*. The signal (*d'*) is the distance between the means of the probability distributions for change in each item in the test display.

In a series of two experiments, human subjects were tested using the change detection task. We first sought to determine if our test parameters would yield similar results to studies published previously. We tested these subjects with similar parameters to what would later be used with monkeys in order to facilitate a cross-species comparison (see Chapters 4 & 5). In the second experiment, four different types of stimuli were tested in order to investigate the contentious role of stimulus type in VSTM.

EXPERIMENT ONE

Methods

Subjects

The six subjects ranged in age from 22-32 (mean age 26.3) and there were five females and one male. The subjects visited the lab for a total of eleven 1-hour sessions.

These sessions were part of a larger study, and only five of these sessions included the test trials presented in this experiment. The subjects were compensated $10 per 1-hour session. All procedures were approved by the University of Texas Health Science Center at Houston Institutional Review Board.

Apparatus

Human subjects were tested in an experimental room with a PC computer. The computer's monitor (17" EIZO) was on a desk in the room and was equipped with an infrared touch-screen (17-inch Unitouch; ELO, Round Rock, TX). The subjects were provided with feedback by two twenty-five watt light bulbs that were mounted on the wall behind the subjects. The green light was illuminated for 1 s following correct responses and the red light was illuminated for 1 s following incorrect responses. The lights were operated by a computer-controlled relay interface (Model PI0-12; Metrabyte, Taunton, MA). Microsoft Visual Basic 6.0 was used to create custom software which created, controlled and recorded experimental sessions. The monitor was controlled by a video card (ATI graphics adaptor).

Stimuli

The stimuli were 6 different colored squares (aqua, blue, green, magenta, red, yellow) and 976 different clip art images. The six colored squares and 12 example clip art images are shown in Figure 2.4. The RGB 24-bit values for the colored squares were aqua – 0, 255, 255; blue – 0, 0, 255; green – 0, 255, 0; magenta – 255, 0, 255; red – 255, 0, 0; yellow – 255, 255, 0. The colored squares were randomly presented in 16 possible locations (defined by points on an invisible 4 by 4 grid) and the clip art items were randomly presented in 20 possible locations (defined by points on two invisible concentric circles). In both cases, the stimuli subtended a visual angle of 1.3 degrees.

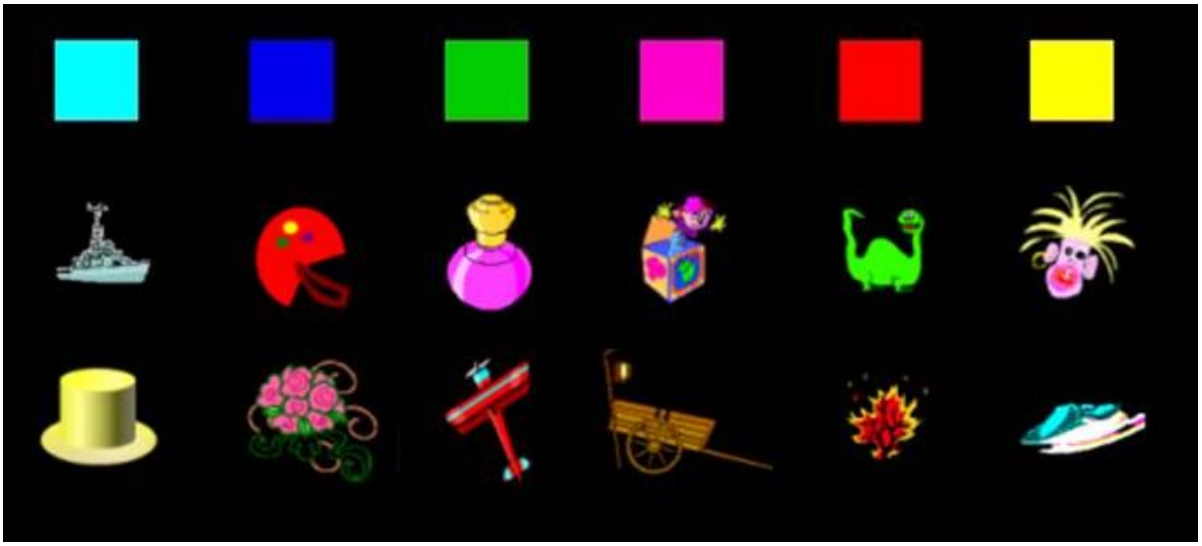


Figure 2.4. Colored Squares and Subset of Clip Art Stimuli. Top Row: Aqua, Blue, Green, Magenta, Red, and Yellow Colored Squares. Middle Row: Battle Cruiser, Football Helmet, Perfume Bottle, Jack-in-the-Box, Dinosaur, and Silly Face Clip Art Objects. Bottom Row: Top Hat, Floral Arrangement, Airplane, Wooden Cart, Burning Bush, and Jet Ski Clip Art Objects.

Test Procedures

The subjects were tested with 150 trials of colored squares (30 trials each of display sizes 2, 4, 6, 8, 10) and 189 trials of clip art (30 trials each of display sizes 2 and 4, and 43 trials each of display size 6, 8, and 10). The viewing time in both conditions was 1000 ms. The delay for colored squares was 900 ms and the delay for clip art was 1000 ms. In both cases the intertrial interval (ITI) was 2000 ms. Colors and clip art were tested in separate sessions, but all display sizes were intermixed within a session. As shown in Figure 2.5, trials began with the presentation of the sample display for 1000-ms. Following the delay (900 or 1000-ms) two stimuli were presented in the test display. One stimulus matched an item presented in the sample display in both identity and location. The other stimulus had changed in identity. Using the touchscreen, subjects were instructed to touch the stimulus that they thought had changed.

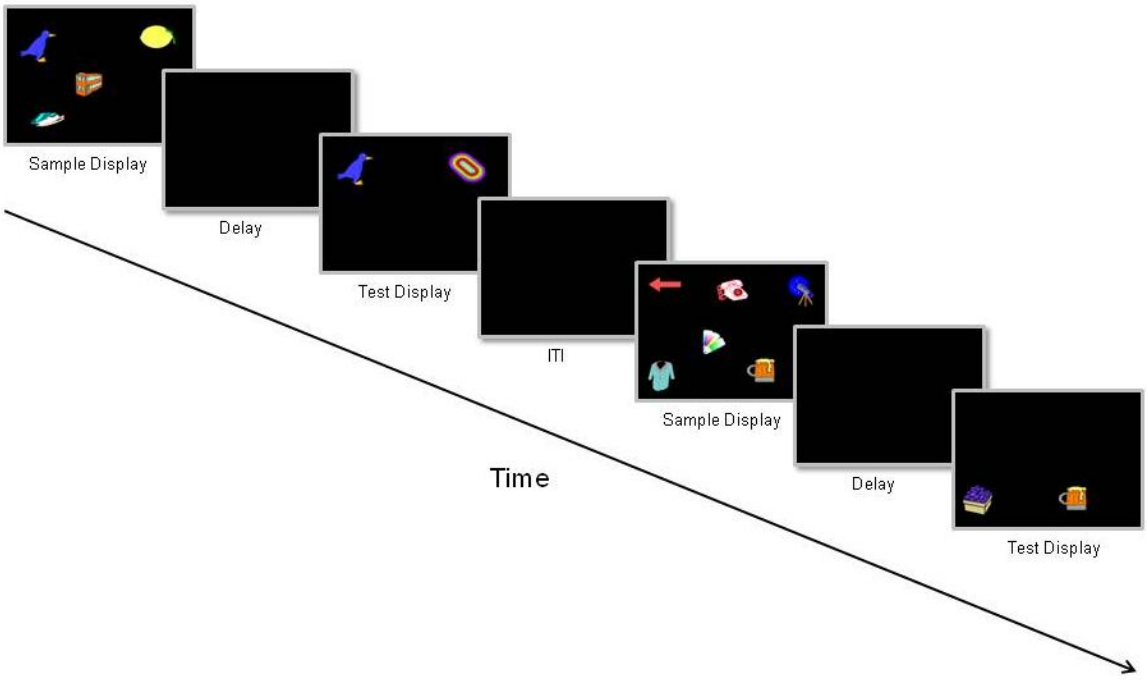


Figure 2.5. Change Detection Task Design. Figure shows two clip art trials. Objects and displays are not to scale with those used in the experiment.

Analysis, Results, & Discussion

As shown in Figure 2.6, performance decreased as display size increased for both colored squares and clip art stimuli. A repeated-measures analysis of variance (ANOVA) of display size × object type showed a significant effect of display size [$F(4,20) = 24.05$, $p < 0.001$]. There was not a significant effect of object type.

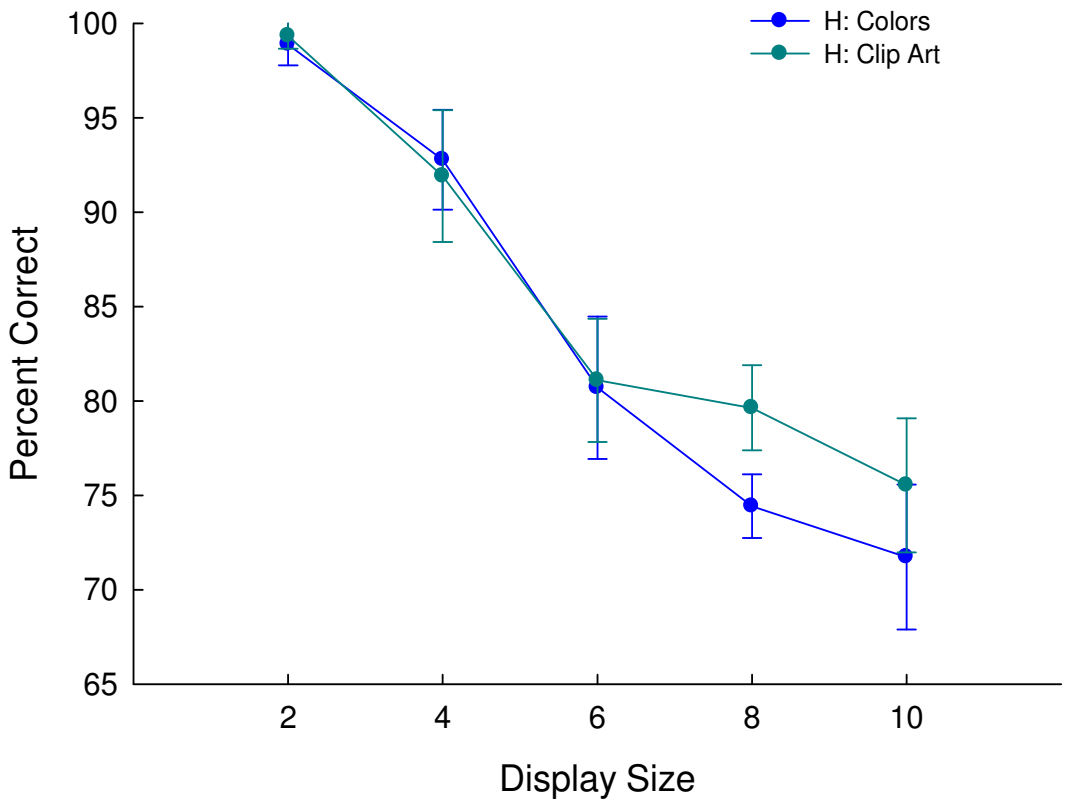


Figure 2.6: Percent Correct by Stimulus Type and Display Size for Colors and Clip Art. Error bars represent standard error of the mean.

Estimating capacity. Percent correct data was used to estimate visual short-term memory capacity using a formula developed by Eng et al. (2005). This formula (Equation 2.1) takes in the empirical accuracy (*A)* and the display size presented (*N*) to solve for capacity (*C*) at that display size.

Equation 2.1:

$$A = \left[\frac{N-C}{N}\right]^2 \times 50\% + \left\{1 - \left[\frac{N-C}{N}\right]^2\right\} \times 100\%$$

As per Eng et al.'s (2005) method, each individual subject's capacity for each stimulus type was estimated by taking the mean of their capacities for each display size (shown in Figure

2.7). Capacity estimates from the display size of two were not included in the mean because two is thought to be less than the average human subject's capacity and would have thus lowered the overall estimate. Mean capacity estimates were 2.46 ± 0.35 for colored squares and 2.78 ± 0.39 for clip art.

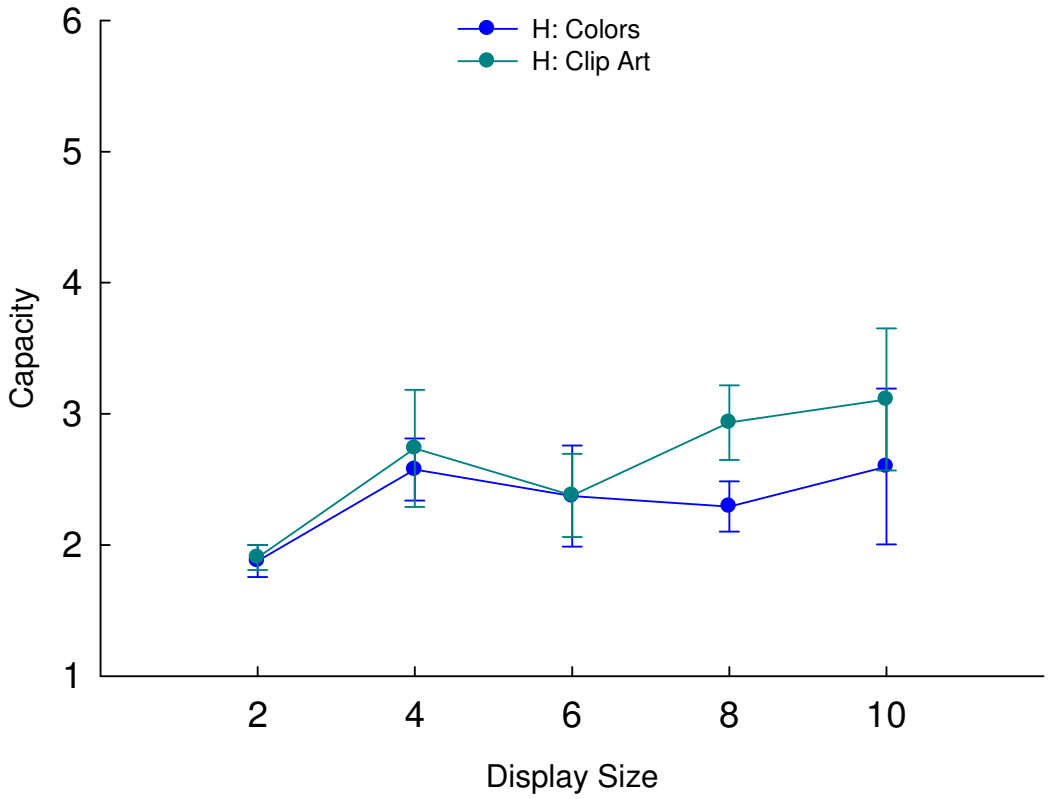


Figure 2.7: Capacity by Stimulus Type and Display Size. Error bars represent standard error of the mean.

Continuous-Resource Model. The results were analyzed according to the continuous-resource model. This model represents memory performance in terms of d' (from signal detection theory). *d'* is a measure of memory sensitivity and the formula to calculate *d'* is shown in Equation 2.2 (Macmillan & Creelman, 2005). *H* is the hit rate, and *FA* is the false alarm rate.

Equation 2.2:

$$d' = \frac{[z(H) - z(FA)]}{\sqrt{2}}$$

The difference of the *z* scores of the hits and false alarms are divided by the square root of 2 because the task is a two-alternative forced-choice task (2AFC) and there are two ways to make a correct response: by remembering that one is the same as the sample display (and picking the other) or by noticing the object that has changed and choosing it. Hits and false alarms were defined based on stimulus location in the test display. Locations were numbered from 1 to 16 (colors) or 20 (clip art) as the locations went from left to right and then down into the row below, and so on. A hit was defined as a correct response to the lower numbered location in the test display. So if test stimuli were displayed in locations 2 and 9 and the stimulus in 2 was the changed object, a correct response to location 2 would constitute a hit. A false alarm was defined as a response to the lower numbered location when that location did not contain the changed item. The definitions of a "hit" and a "false alarm" are arbitrary but equivalent to the obverse. *d'* values for each stimulus type and display size are plotted in Figure 2.8.

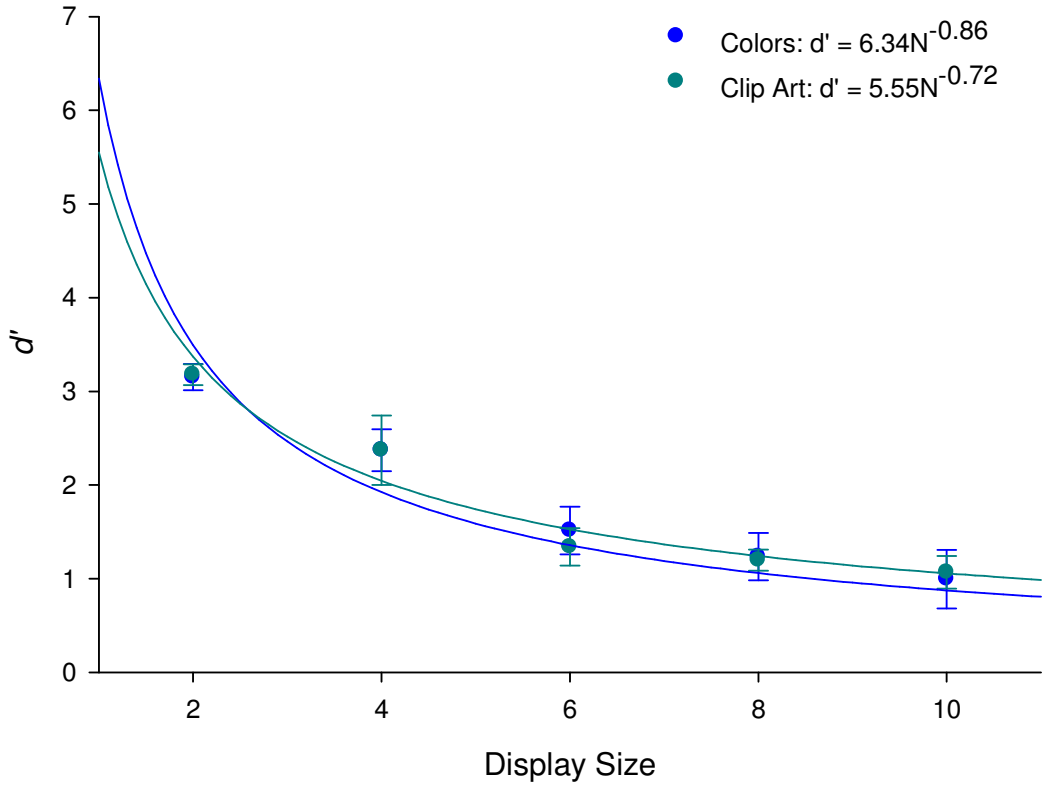


Figure 2.8: Power Law Fits for *d'* for Colors and Clip Art.

The *d'* values for each stimulus type were fit with power law functions using Microsoft Excel 2007. First, individual subjects' *d'* values were plotted. These plots were fit with power law functions using the trendline function in Excel. The mean power law functions (displayed in Figure 2.8.) were produced by taking the mean of the *d'* values produced by each individual subject's best fit power law function (by solving for each *d'* using the known display sizes), and then fitting a power law to those means, again using the trendline function in Excel. The power law functions were found to provide a good fit to the *d'* values (r^2 values were 0.75 for colors and 0.70 for clip art). r^2 values were obtained by conducting a regression analysis comparing all subjects' combined empirical *d'* values to their combined predicted *d'* values. Predicted *d'* values were obtained from individual subjects' power law fits. The r^2 values of

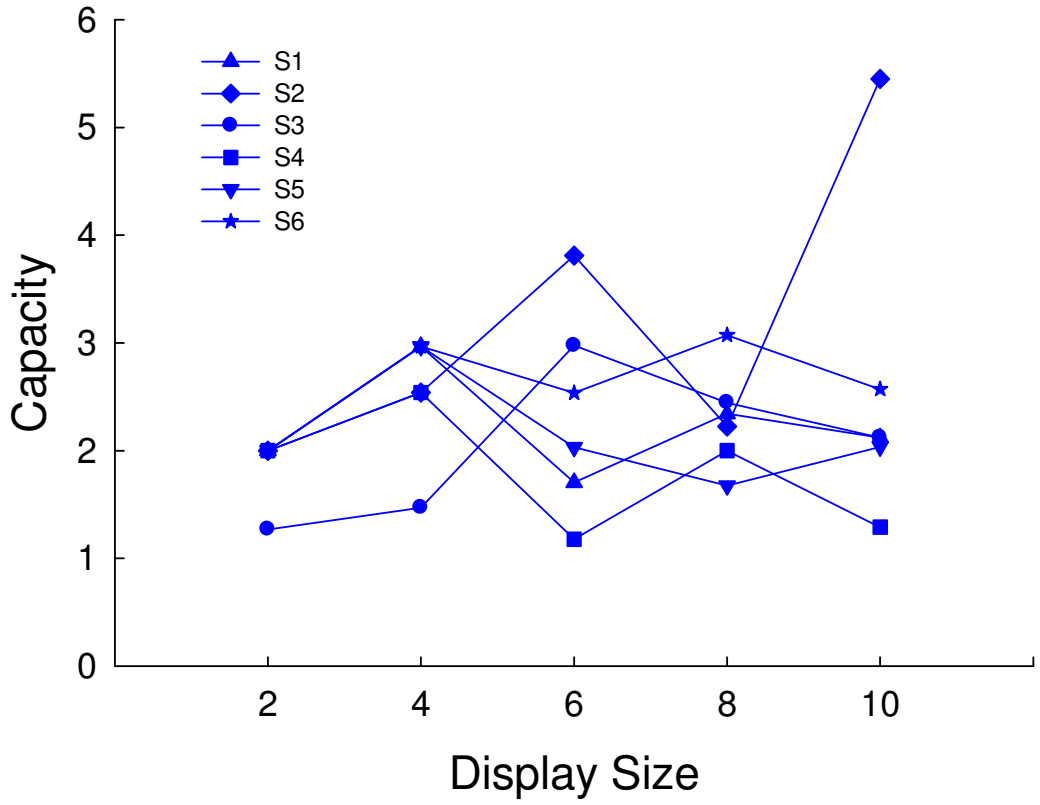
0.75 (colors) and 0.70 (clip art) were found to be highly significant [Colors: $F(1,28) = 81.97$, $p < 0.0001$; Clip Art: $F(1,28) = 65.42$, $p < 0.0001$].

Not surprisingly, the results of Experiment 1 demonstrate that visual short-term memory performance falls as a function of display size. Memory is worse when there are more items to remember. The capacity measures found here are consistent with values published previously using the same procedures. For instance, Eng et al. (2005) found capacity measures of 2.4-2.5 for colored squares, values that are essentially identical to our value of 2.46 ± 0.35. However, the capacity measures obtained in this experiment are somewhat lower than the value popularized in the literature of 4±1 (e.g. Cowan, 2001; Cowan, 2005). However, the procedures of Cowan and others are slightly different than the procedures employed here. These prior studies have often used a "change/no change" procedure where the subject is presented with the same number of items in both the sample and test display. In half the trials a change occurs, and in the other half there is no change (Alvarez & Cavanagh, 2004). The subject is asked to judge whether or not a change has occurred. This procedure is potentially easier than the two alternative forced choice task used here because the presence of all the sample items in the test display may provide contextual cues which enhance VSTM performance.

While it is important to estimate capacity at multiple display sizes and take the mean to get a true estimate of an individual subject's capacity, there needs to be some amount of consistency across capacity measures for varying display sizes. Interestingly, as shown in Figure 2.9 there was considerable within-subject variability in capacity estimates across display sizes. For example, S6's capacity estimates in the clip art condition ranged from 1.65 (in the four item display) to 5.29 (in the ten item display). From the perspective of the fixed capacity's hallmark, the magic number 4 ± 1, it is reasonable to expect that capacity estimates across display sizes for an individual subject should fall within one standard deviation of that subject's mean capacity estimate, because for the majority of the population, capacity

estimates supposedly fall within one standard deviation of the magic number four. Thus, taking the example of S6, the subject's mean capacity estimate for clip art was 2.94, such that a capacity estimate of 5.29 in the ten item display well exceeds one standard deviation (2.94+1.53 = 4.37). Such variability in capacity measures is difficult to reconcile with the construct of a capacity-limited storage mechanism. If visual stimuli are truly stored in a "slot-like storage system" then why should the capacity of that storage system vary so widely? The variance cannot be attributed to the different display sizes, because the equation used to compute capacity (Equation 2.1) takes display size in to account.

A



B

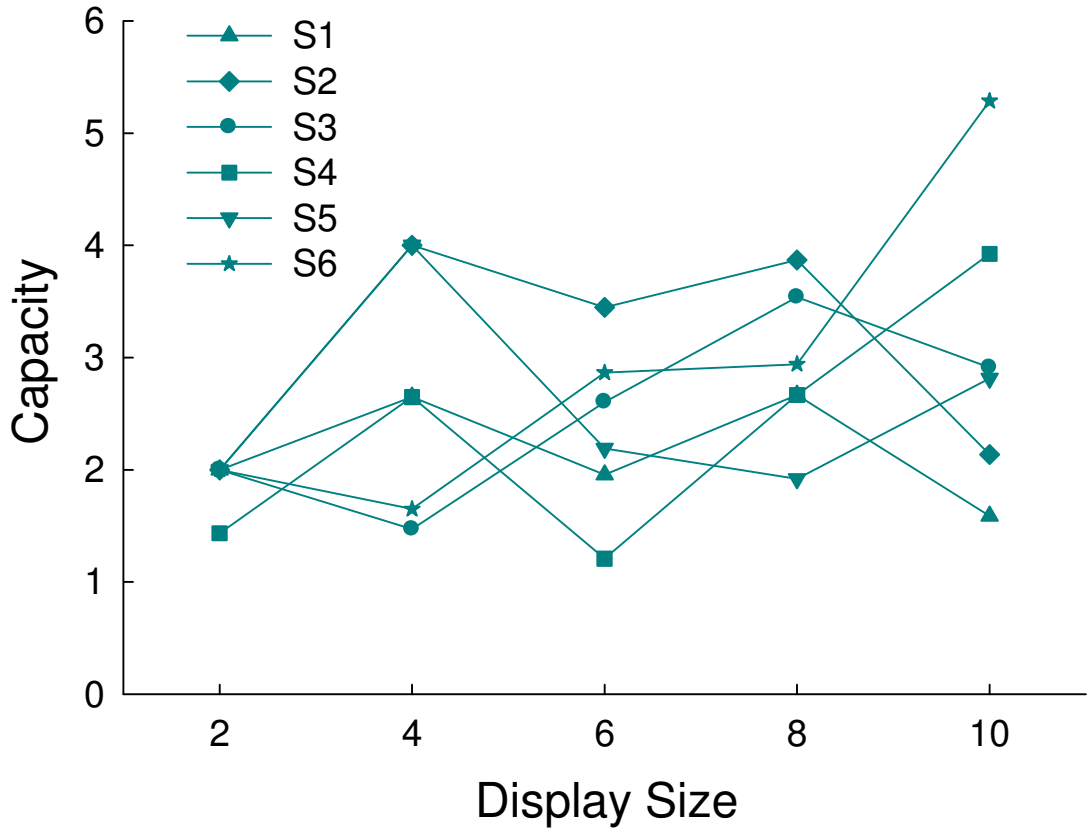


Figure 2.9: Individual Subject Variability in Capacity Measures for A) Colors and B) Clip Art.

Another concern regarding the fixed-capacity model is that the model predicts that percent correct at display sizes less than capacity (e.g., 2) should be perfect (100%). If the subject has more than enough slots than needed to accommodate the items in the display then they should perform with 100% accuracy. However, two subjects had less than perfect

performance in the easiest two-item display size condition. S3 was 93.3% accurate (capacity of 1.27) for colors and S4 was 96% accurate (capacity of 1.43) for clip art.

The continuous-resource model provides a good fit to the results of Experiment 1. For both the clip art and the colored squares, the r^2 values were high, showing that power law functions provided a good fit to the group means of *d'* values. In addition, the mean power of the power law functions (0.79 ± 0.07) is very similar to the power value (0.74 ± 0.06) reportedly recently by Bays & Husain (2008). Thus, not only does the continuous-resource model provide a good fit to the data, but the fit is strikingly similar to one reported by another group providing some converging evidence in favor of the continuous-resource model of VSTM.

EXPERIMENT 2

Methods

Subjects

Seven subjects ranging in age from 23 to 32 (mean age 26) participated in Experiment 2. There were five females and two males. Five subjects had also participated in Experiment 1. These subjects participated in a total of 11, 1-hour sessions as part of a larger study. The other two subjects participated in 8, 1-hour sessions, again as part of a larger study. The results presented here are from two 1-hour sessions which all seven subjects completed in their entirety.

Apparatus

The apparatus was the same as described in Experiment 1.

Stimuli

In Experiment 2, subjects were tested with four types of stimuli (three of which are shown in Figure 2.10); Clip Art (shown in Figure 2.4), Kanji characters, Kaleidoscope images, and Snodgrass black and white line drawings. The Kanji characters and Snodgrass line drawings were drawn from a set of 256 stimuli and the clip art and kaleidoscope images were

drawn from a set of 976. The clip art sessions included two of the same sessions tested in Experiment 1. The smaller display sizes (two and four) tested in Experiment 1 were tested in a separate session, and those results will not be included in this experiment. The stimuli were presented in twenty possible locations (defined by points on two invisible concentric circles) and subtended a visual angle of 1.3 degrees.

Figure 2.10: Experiment 2 Stimuli. Top Row: Kanji Characters. Middle Row: Kaleidoscope Images. Bottom Row: Snodgrass Line Drawings (Envelope, Stool, Television, Axe, Mountain, and Necklace).

Test Procedures

Over the course of two sessions (conducted on different days), subjects were tested with 90 trials of each stimulus type (30 each of display sizes 6, 8, 10). The viewing time and delay were both 1000 ms and the intertrial interval was 2000 ms. All display sizes and stimulus types were intermixed within a session.

Results, & Discussion

As shown in Figure 2.11, performance decreased as display size increased, for Kanji characters, kaleidoscopes, Snodgrass line drawings, and clip art. A repeated measures ANOVA of display size × stimulus type revealed a significant effect of both display size

[$F(2,12) = 18.5$, $p < 0.001$] and stimulus type [$F(3, 18) = 16.412$, $p < 0.001$]. There was not a significant interaction. Subjects performed best with Snodgrass line drawings and clip art and performed worst with Kanji characters and kaleidoscopes.

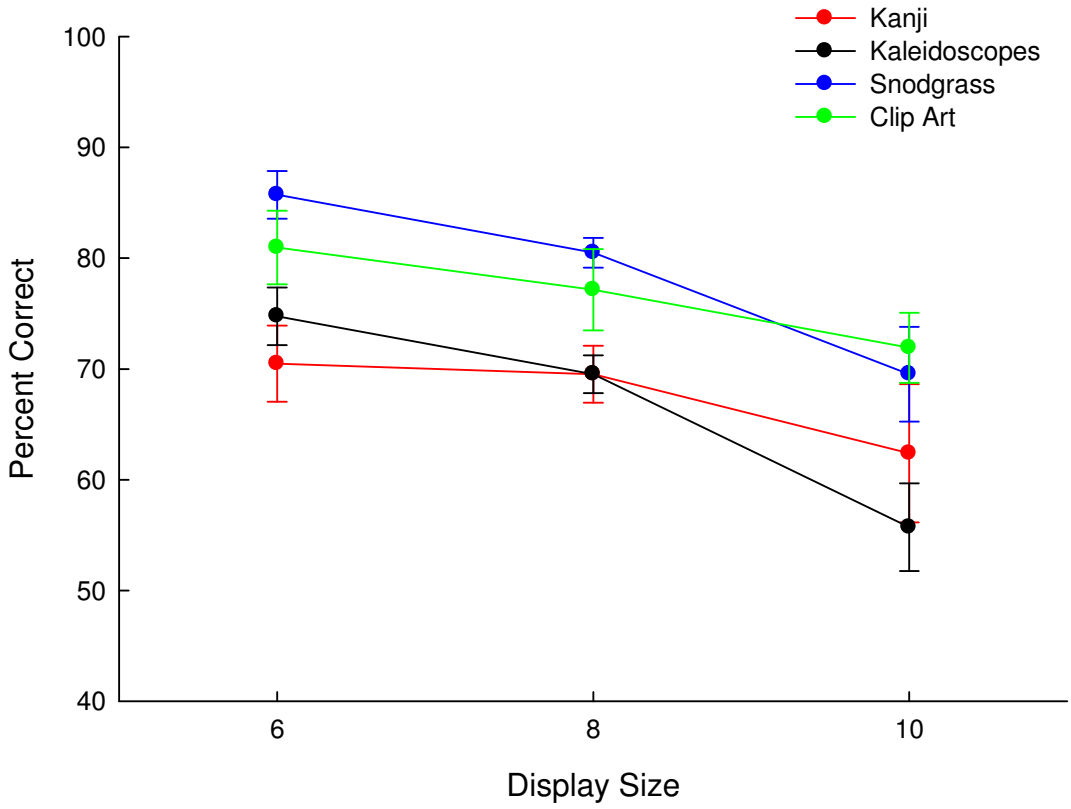


Figure 2.11: Percent Correct by Stimulus Type and Display Size for Kanji, Kaleidoscopes, Snodgrass, and Clip Art. Error bars represent standard error of the mean.

Estimating capacity. Capacity measures were calculated using Equation 2.1. Mean capacity limits for each display size and stimulus type are shown in Figure 2.12. As with accuracy, capacity measures were higher for Snodgrass line drawings (mean of 2.73 ± 0.22) and clip art (2.55 ± 0.09) than for Kanji (1.59 ± 0.10) and kaleidoscopes (1.39 ± 0.37).

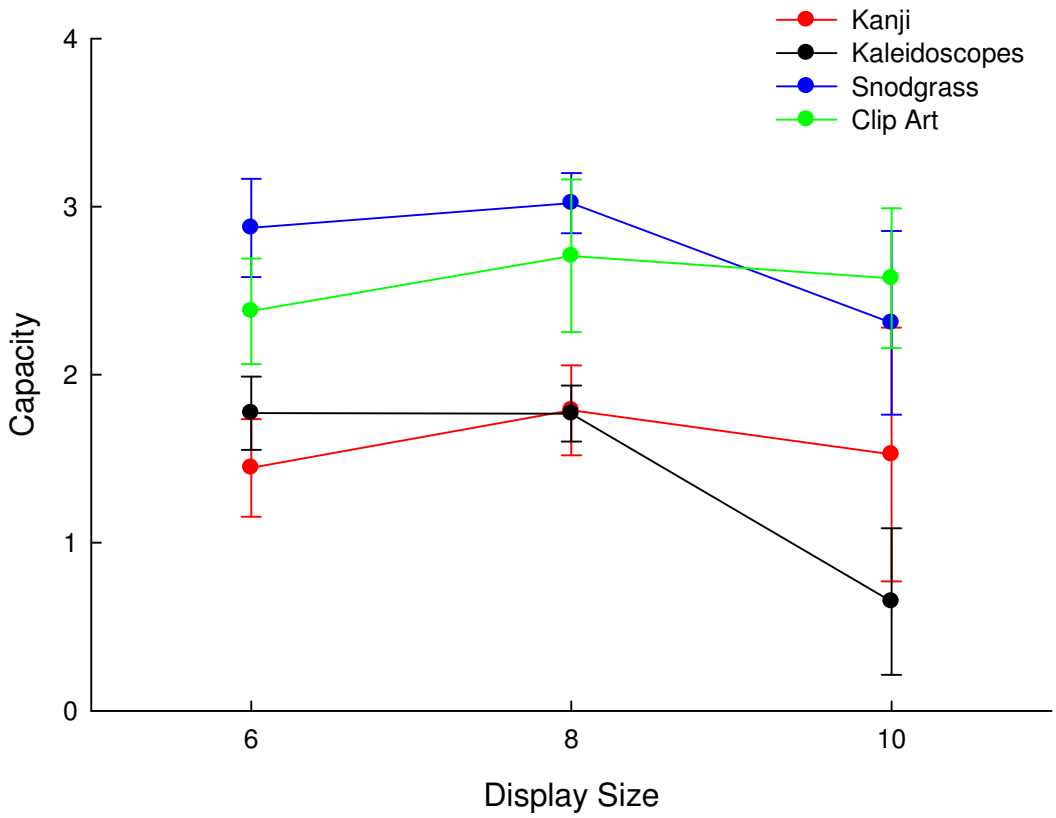


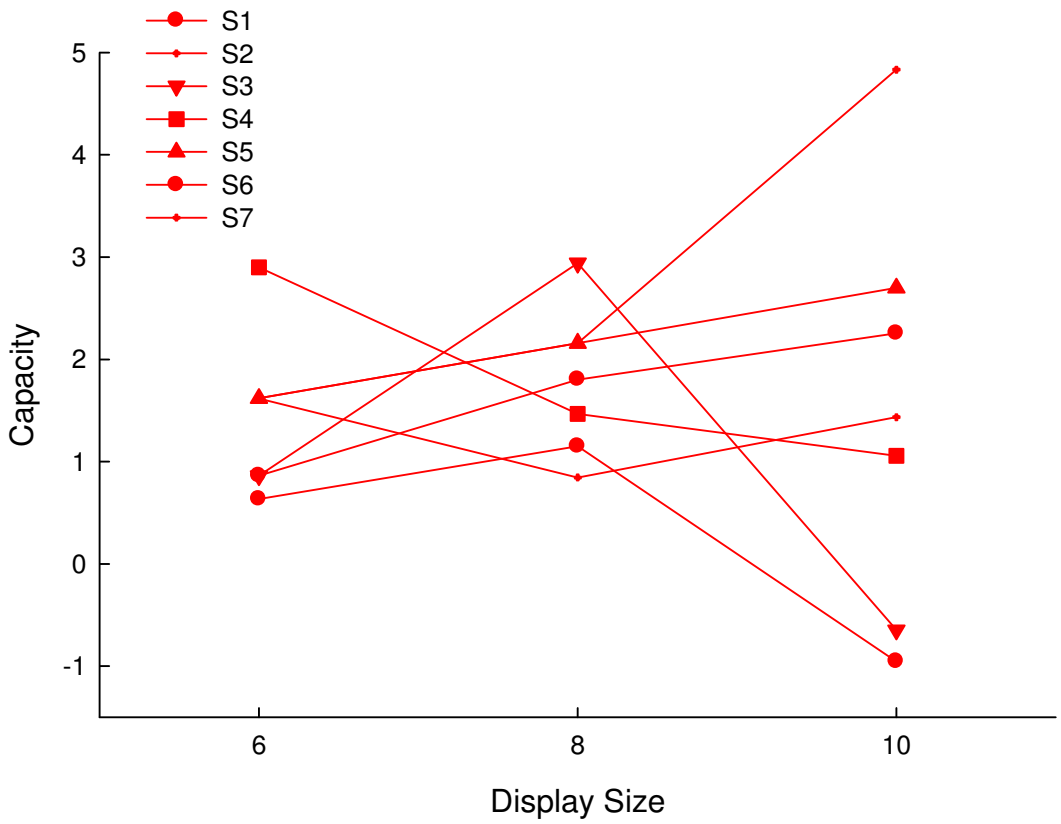
Figure 2.12: Capacity Measures by Stimulus Type and Display Size. Error Bars represent Standard Error of the Mean.

A repeated measures ANOVA of display size × stimulus type revealed a significant effect of stimulus type [$F(3, 18) = 12.93$, $p < 0.001$]. As predicted by the fixed-capacity model, there was no significant effect of display size, because capacity measures are supposed to be independent of display size.

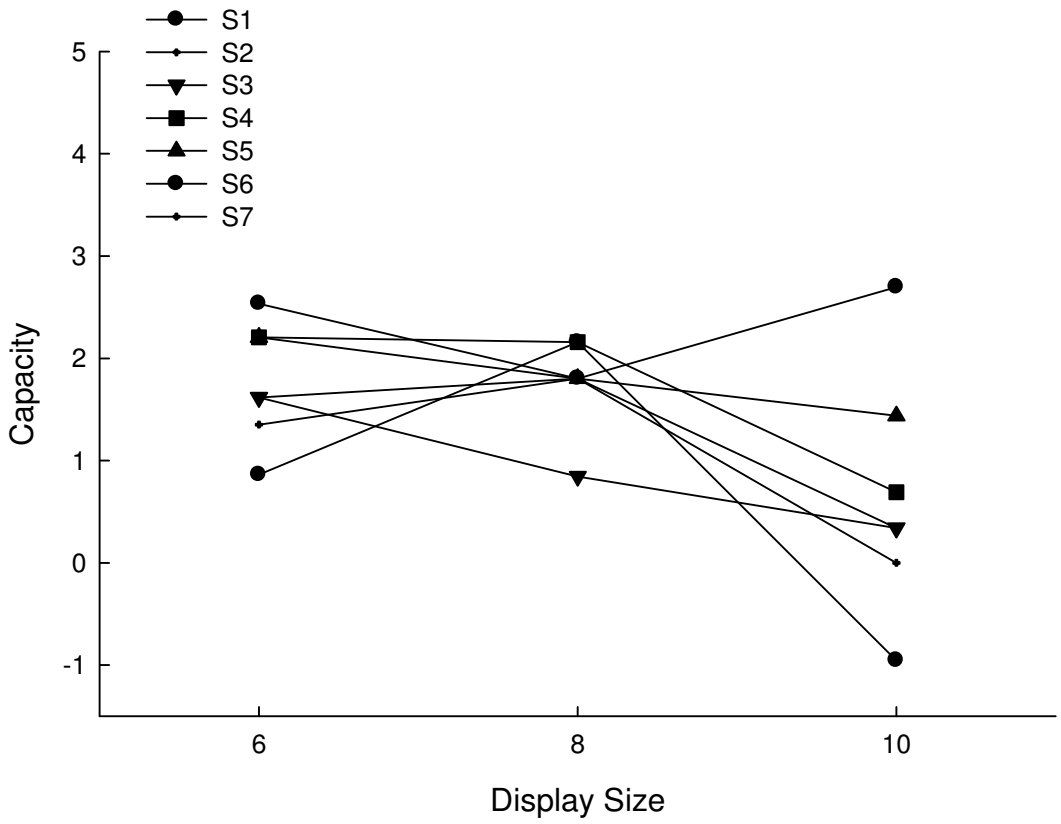
However, once again, as in Experiment 1, there was a good deal of within subject variability in capacities for a given stimulus across the three display sizes tested. As shown in Figure 2.13, some subjects showed somewhat stable performance across display sizes, whereas others were widely variable. For instance, when tested with Kanji characters, S2's capacity measures range from 1.62 in the 6 item display size to 4.83 in the 10 item display size. Worse yet, is the finding of negative capacities in the 10 item display size for S3 and S6.

Performance of 50% correct (chance performance) at any display size yields a capacity limit of 0. S3 and S6 were 43.33 and 40% correct respectively at display size 10, resulting in negative capacities. Finding a capacity limit of 0 or less is conceptually implausible. Formulas used to compute capacity are supposed to take display size into account. The notion of a fixed-capacity originates from the desire to understand the limitations of VSTM, and in order to do so, one must challenge subjects by testing them in conditions which push the limits of their VSTM abilities. It is a very unsatisfactory finding that performance by some subjects at large display sizes is indicative of them having stored nothing in VSTM (as suggested by a capacity measure ≤ 0).

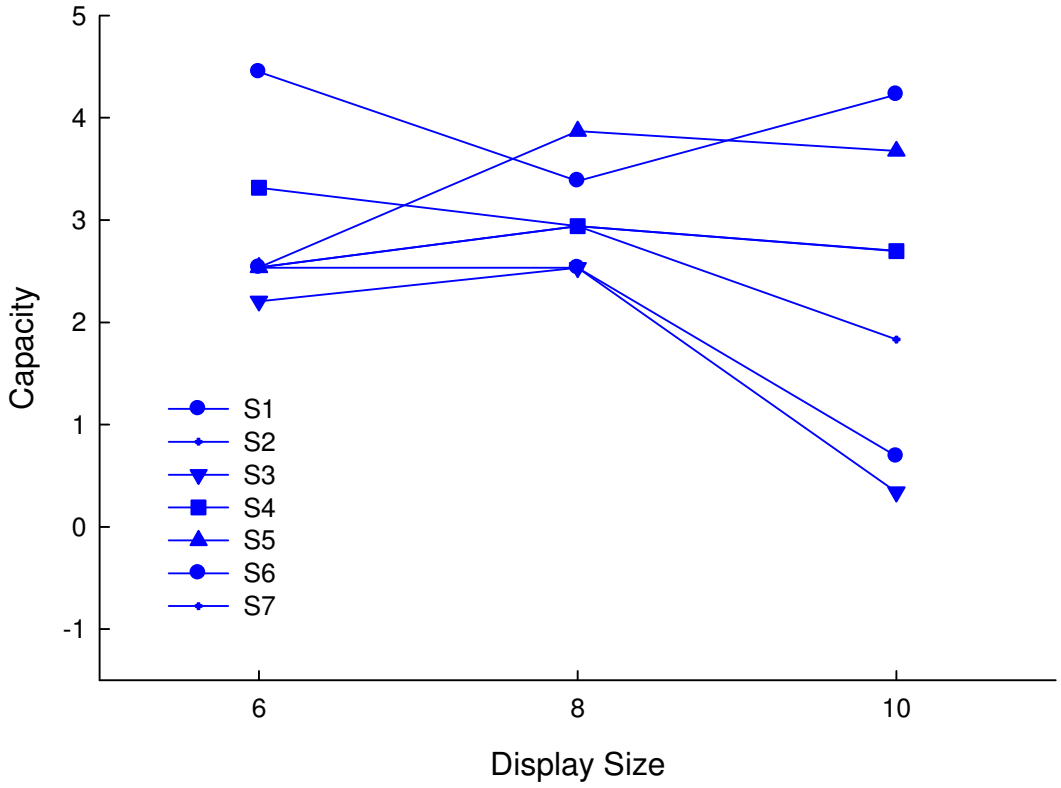
A



B



C



D

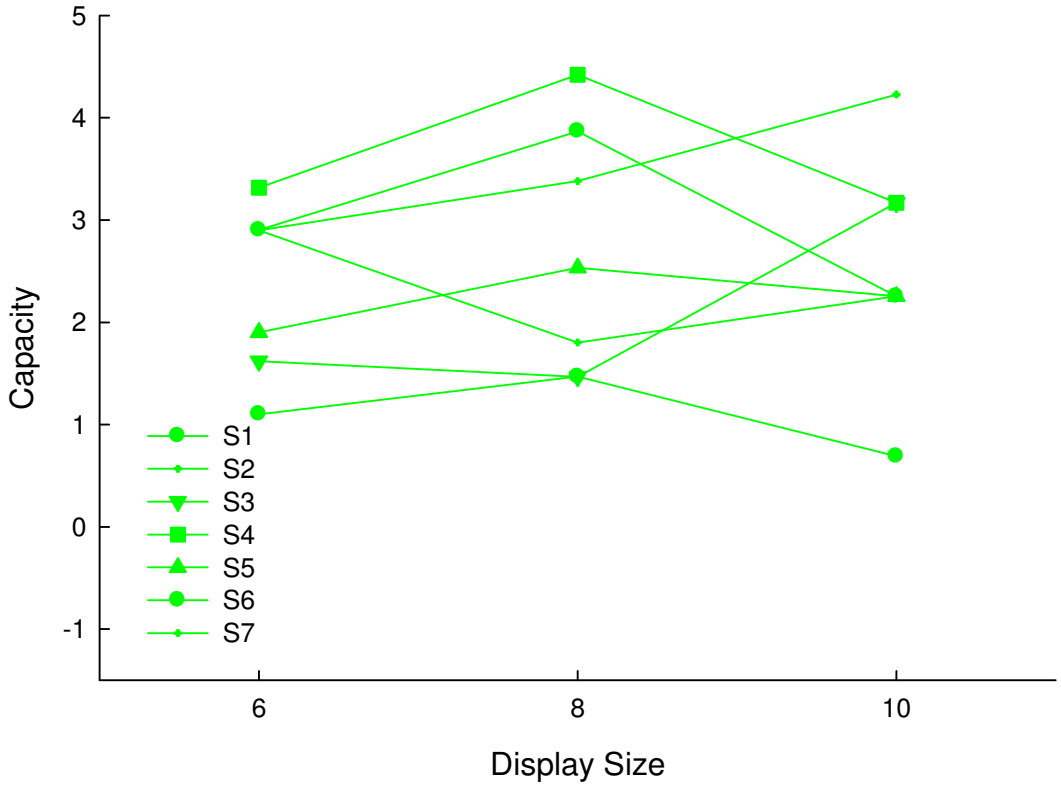


Figure 2.13: Individual Subject Variability in Capacity Estimates for A) Kanji Characters, B) Kaleidoscope Images, C) Snodgrass Line Drawings, and D) Clip Art.

A main assumption of the fixed-capacity model seems to fail. The assumption is that capacity measures should be stable across display sizes. Of the seven subjects tested in Experiment 2, six showed capacity measures with differences of 1.5 or greater across display sizes for at least one stimulus type. With capacity measures on average being 2.06 (mean of the four stimulus types), a difference in capacity of 1.5 or greater represents a substantial change in capacity of 75% or greater for a single stimulus type. Even taking the perspective of a flexible fixed-capacity model (which varies by stimulus type) one cannot reconcile 75% changes in capacity across display sizes for any given stimulus type with the notion of a capacity limited slot-like storage system. Such inconsistency questions the very existence of

such a system. Taking the example of S2 with Kanji characters, why is it that when the subject has six kanji characters to remember there are "slots" for only 1.62 characters, whereas with ten display items size, there are "slots" for 4 or more? Small disparities in capacity measures could be attributed to sampling noise, but huge changes such as S2's near tripling in capacity (from 1.62 to 4.83) are incompatible with a slot-like storage system of fixed capacity.

Our findings also reject the more rigid interpretation of the fixed-capacity model, the idea that capacity should not depend on stimulus type / visual complexity, as we found clear differences in capacity measures across the four stimulus types tested. Eng et al. (2005) would argue that with longer viewing times, these differences should wash out, and capacity measures would equate. Although we only tested one viewing time (1 second), Eng et al.'s claim that capacity measures level out with extended viewing times tells us very little about how memory works. They allowed their subjects to view the sample display for as long as they wanted and then concluded that capacity measures were equal after adequate time viewing the stimuli. They proposed that increasing the viewing time allowed subjects sufficient time to adequately perceive the objects (thereby improving performance), but since they do not report the durations that subjects chose to view the items it is unclear whether the increase in viewing time was necessary for perception, or rather more likely provided additional study time and/or a reduction in proactive interference from earlier trials leading to enhanced performance with difficult stimuli. Furthermore, in order to directly compare memory for different types of stimuli, all other variables must be held constant. Thus, our finding of capacity (and performance differences) across stimulus types is likely a more valid finding given that our experiment was more tightly controlled.

Continuous-resource model. As described in Experiment 1, *d'* values were computed based on the accuracy results presented in Figure 2.11. The *d'* values are plotted in Figure 2.14. As shown in the figure, the values were all well fit by power law functions, with high r^2 values (0.87 for Kanji, 0.85 for kaleidoscopes, 0.87 for Snodgrass, and 0.99 for Clip Art), and

highly significant *p* values [Kanji: $F(1,19) = 38.27$, $p < 0.0001$; Kaleidoscopes: $F(1,19) = 35.66$, $p < 0.0001$; Snodgrass: $F(1,19) = 29.77$, $p < 0.0001$; Clip Art: $F(1,19) = 110.23$, $p < 0.0001$].

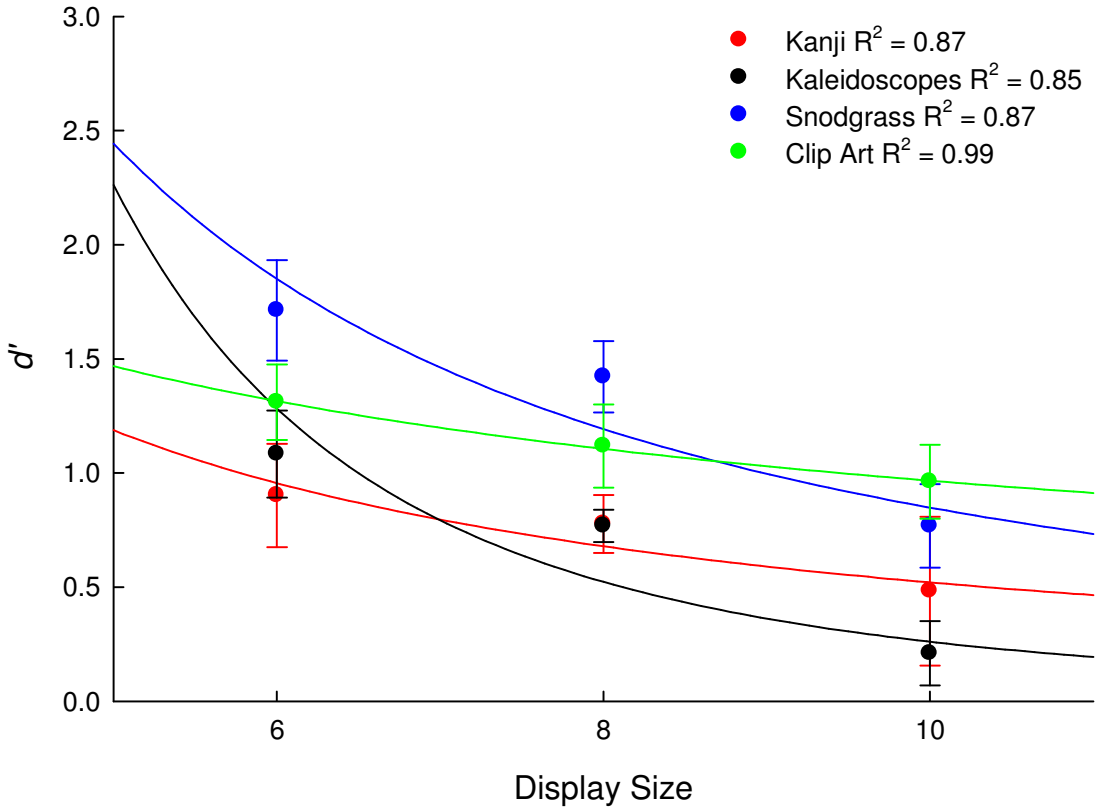


Figure 2.14. Power Law Fits for *d'* values for Kanji, Kaleidoscopes, Snodgrass Line Drawings, and Clip Art. Error bars represent standard error of the mean.

Thus it seems that the continuous-resource model may provide a better fit to our results than does the fixed-capacity model. The relatively simple predictions of the continuous- resource model, that *d'* should fall as a function of display size, and that the decline in *d'* should be well fit by a power law function were met for all four stimulus types tested.

One advantage of the continuous-resource model over the fixed-capacity model is its ties to the neural basis of memory. While it is difficult to posit a neural mechanism that could give rise to a slot-like storage system, the continuous-resource model's predictions of noisy

representations and probabilistic decisions are very much in line with what is known about neurobiological computations in general (e.g., Bays & Husain, 2008; Beck et al., 2008; Ma et al., 2006, Ma et al., 2008). What remains to be seen is whether or not the continuous-resource model can account for VSTM in another species, the rhesus monkey.

CHAPTER 3: TRAINING RHESUS MONKEYS TO PERFORM CHANGE DETECTION

Introduction

The purpose of this experiment was to train and test rhesus monkeys in the change detection task. In order to hasten the monkeys' acquisition of the task, they were trained in the simplest possible version of the task. During their initial training, the monkeys had a minimal number of stimuli to remember (two), a long time to view and encode the stimuli (5000-ms) and a very brief retention delay (50-ms).

As the task was originally established for use with humans, it was important to demonstrate that rhesus monkeys could learn the task, and perform in a way analogous to human subjects. Humans can be instructed to look for all types of changes prior to performing the task. Monkeys, however, must learn the rules through the contingencies of reinforcement. Specifically, it was necessary to demonstrate that they understood a concept of "change" as evidenced by an ability to detect changes that they were not explicitly trained to detect. In many other behavioral tasks with animals, they are found to learn a strategy which is very specific to their training conditions. One example is the finding of item-specific learning strategies in pigeons and monkeys trained to perform a same/different task (e.g. Wright & Katz, 2006, Elmore et al., 2009). Instead of learning the concept of "same" and "different", the animals memorized correct responses to individual pairs of stimuli. Because a similar response memorization strategy could occur with monkeys trained in change detection, we tested monkeys with novel colors and shapes, and novel types of change including shape, location, and size changes to see if their learning of change would be general.

In addition, the monkeys were trained with very short delays (50-ms) to enhance acquisition, therefore tests were conducted to demonstrate that they were performing the task using mnemonic processing as opposed to an attentional capture mechanism (e.g., Cusack, 2009; Pashler, 1988; Yantis, 1993). The bottom-up process of attentional capture can occur when a stimulus abruptly changes or from the sudden onset or offset of a stimulus. An internal mechanism guides attention to this locus of change. If the monkeys' performance could be

explained by attentional capture, then they would unlikely be using short-term memory to encode the stimuli. For this test, the monkeys were abruptly (i.e., without gradually increasing the delay through training) transferred to sessions composed of trials with variable delays ranging from 100 to 6400-ms. The delays which ranged from 800-ms to 6400-ms are considerably beyond the time scale of attentional capture (as studied in humans). If an attentional capture mechanism was responsible for the monkeys' performance, then performance should have fallen abruptly to chance at delays of 800-ms or greater.

Methods

Animals

Two adult males rhesus monkeys (*Macaca mulatta)*, M1 and M2, were the subjects. M1 and M2 were eight and twelve years old respectively at the start of the study. Both monkeys had prior experience in same/different and list-memory tasks, however these tasks were conducted in a different chamber with different stimuli, different display configurations and different response templates. The monkeys were tested five days per week for sessions that lasted a maximum of two hours per day. They were not fed or given water in the morning before their session, but were fed a ration of primate chow and water in their home cages after the daily session. On weekends (non-testing days), the monkeys were provided supplemental fruits and vegetables. Animal procedures were in compliance with the National Institutes of Health guidelines and were approved by the University of Texas Health Science Center at Houston's Institutional Animal Care and Use Committee.

Apparatus

Chambers. Custom-made aluminum test chambers were used to test the monkeys. The chambers were 47.5 cm wide × 53.13 cm deep × 66.25 cm high. Restraints were not used; the monkeys were able to move freely within the confines of the test chamber. A sound machine (Homedics, Commerce Township, MI) located outside of the chamber was used to

produce white noise to mask extraneous noise. A 17" computer monitor (EIZO) equipped with an infrared touchscreen (Unitouch, ELO, Round Rock, TX) was fitted in the back wall of the chamber 30 cm above the chamber floor on which the monkeys sat. The touchscreen was used to detect touch responses to the computer monitor. On the left side of the back wall of the chamber, 14 cm below the touchscreen, was a pellet cup (5.6 cm in diameter, 2.5 cm deep) which received delivery of banana pellets (Bio-Serv, 300-mg, Frenchtown, NJ) from a pellet dispenser (Gerbrands, G5-120, Arlington, MA) located outside of the chamber. Cherry-Koolaid was dispensed via plastic tubing to a metal spout located 8 cm below the touch screen on the right side of the chamber's back wall.

Stimuli and Display Parameters. The monkeys were initially trained with 4 different colored circle stimuli that were 4 cm in diameter. The RGB 24-bit values for these stimuli were: Red – 255, 0, 0; Aqua – 0, 255, 255; Yellow – 255, 255, 0; and Purple – 180, 0, 255. Later, during testing, 4 additional colors were added, whose RGB 24-bit values were: Blue – 0, 0, 255; Green – 0, 255, 0; Magenta – 255, 0, 255; and Orange – 255, 128, 0. All eight colors are shown in Figure 3.1.

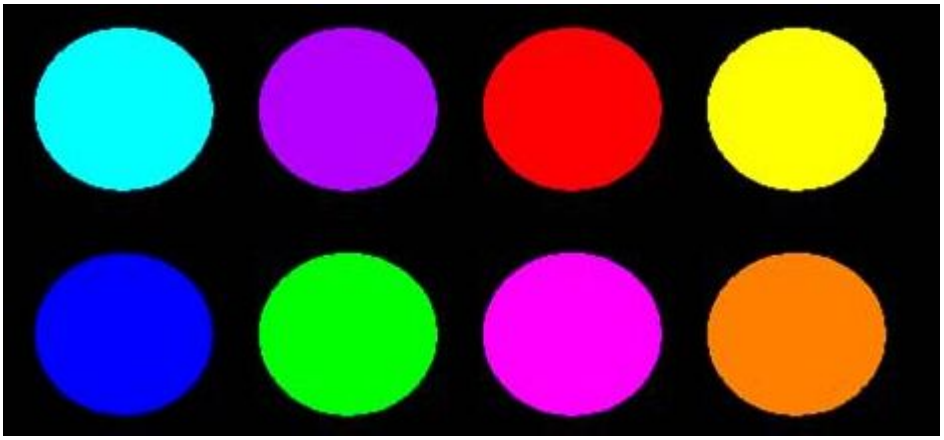


Figure 3.1: Colored Circle Stimuli. Top row: training colors. Bottom row: test colors.

Novel shapes used for testing included butterfly, club, heart, pentagon, rectangle, star, and triangle. The stimuli were presented within an invisible 4 × 4 matrix on the computer

monitor, which was aligned to a clear Plexiglas template placed in front of the monitor. The template had 16-circular 4-cm cutouts that the monkeys were required to reach through in order to touch the stimuli.

Experimental Control. Microsoft Visual Basic 6.0 was used to create, control and record experimental sessions. An ATI graphics adaptor video card was used to control the monitor. The green light, pellet dispenser, and juice system were controlled and operated by a computer-controlled relay interface (Model PI0-12; Metrabyte, Taunton, MA).

Training Procedures. Because both monkeys were naïve to the change detection procedures, they began with a pretraining procedure designed to acclimate them to the new chamber and template. They completed 96-trial pretraining sessions in which achromatic circular stimuli were randomly presented (1 circle per trial) for 3 seconds in the 16 positions on the invisible by 4×4 grid. On each trial, the circle would change from white to grey or from grey to white after a 50-ms delay. Monkeys were provided banana pellets or Cherry Kool-aid for making a touch response to the circle after it had changed. Pretraining trials were separated by a 15-second intertrial interval (ITI). During the ITI, two green (25 watt) light bulbs located outside of the chamber provided illumination through a small gap between the touchscreen and the monitor.

M1 was averse to the new chamber and did not respond reliably until he had completed 10 sessions of pretraining. During this time, extensive hand shaping (reinforcing successive approximations to the required response, controlled by the experimenter) was required to encourage responding. M1 also completed three sessions of the familiar same/different task in the new chamber (with the template removed) which improved his responding. In the last phase of pretraining, the achromatic circles were replaced with the four training colors (red, yellow, aqua, purple). M1 then completed three 96-trial sessions wherein a single randomly selected color was presented for 3000 ms and then changed to one of the three other colors after the 50ms delay. The first response following the change was

reinforced. The other monkey, M2, was more willing to make responses in the new chamber. He only required two pretraining sessions, one with achromatic circles and one with colored circles before starting change-detection training.

Change detection training began immediately following pretraining. Figure 3.2 shows two examples of change-detection trials. The trials commenced with a 5000-ms presentation of two different colored circles in two randomly selected positions on the invisible 4×4 matrix (sample display). The trials were counterbalanced such that there was an equal likelihood of each color appearing as a sample stimulus or changed-to color. Following a 50-ms delay (with a blank screen), the two circles reappeared, but one had changed in color.

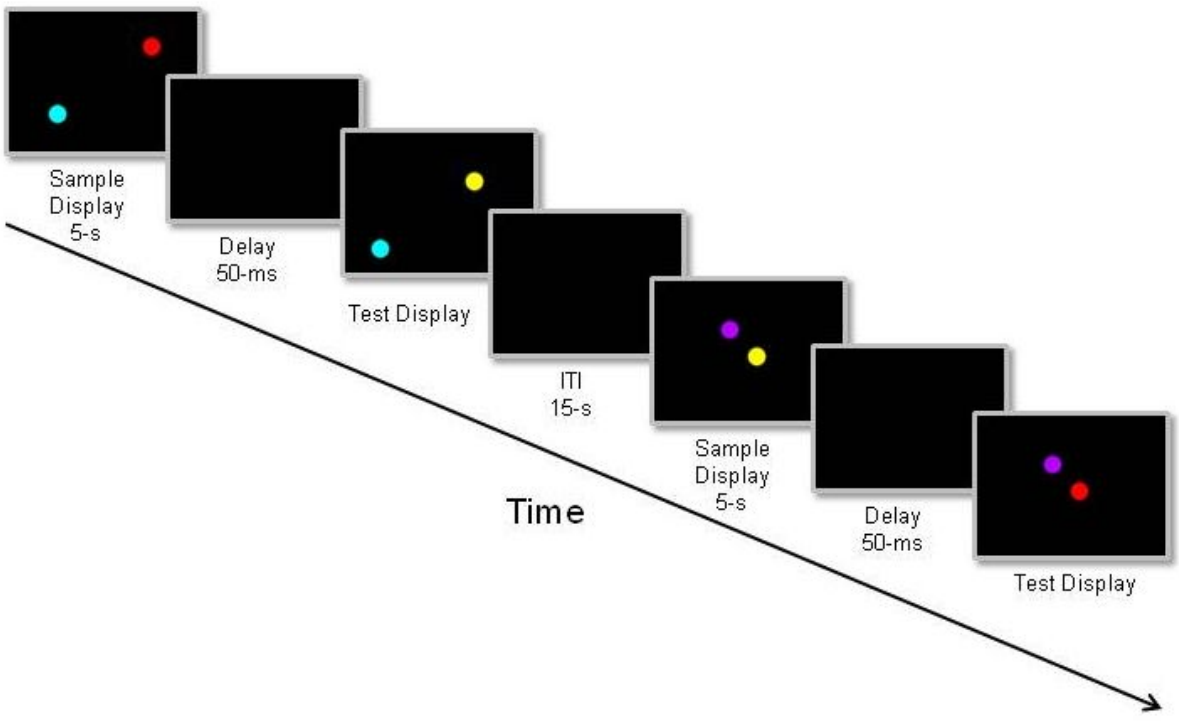


Figure 3.2: Progression of the Change Detection Task. Two trials are shown.

The monkeys' task was to touch the circle that had changed color and banana pellet and Cherry Kool-aid reinforcers were provided pseudorandomly following correct responses. The pellet and juice ratios were determined based on the monkeys' preferences, so M1 received juice on 70% of correct trials and pellets on the other 30% whereas M2 received 50% juice and 50% pellets. All other experimental details were the same as described for pretraining including session length (96 trials). The monkeys were allowed two hours to complete a session in a given day. If they did not complete the session within two hours, it was continued the following day, but this rarely occurred.

During the course of acquisition, several procedural manipulations were used in an attempt to hasten acquisition of the task. Both monkeys had a shorter 0-ms delay for a maximum of 14 days during training. M2 experienced this short delay early in acquisition period whereas M1 had the 0-s delay introduced in the middle of his training. In addition, M2 had a shorter 5-s ITI during the first 48 training sessions. Both monkeys had a correction procedure, where incorrect trials were repeated until a correct response was made. The correction procedure was started on the third session of training and continued until the monkeys achieved ≥ 80% correct on a session. Sessions with 0-ms delays were not counted towards this 80% criterion. After the criterion was met, the correction procedure was removed and training continued until the monkeys again performed 80% correct on a session, at which point they met criterion for acquisition, and could begin testing.

Tests

Color transfer. Once criterion for acquisition had been met, monkeys were tested for transfer to novel colors over the course of six consecutive test sessions. Each session was 96 trials in length with twelve test trials composed from a set of four novel colors (blue, green, magenta, and orange) randomly dispersed throughout. The other 84 trials were baseline trials, composed from the training colors. On test trials, each of the novel color stimuli appeared as one of the two circles in the sample array on six trials and appeared as the

change-to stimulus in three of the six test trials where that color was not in the sample array. Correct responses on both test and baseline trials were always reinforced.

Variable delay testing. After the novel color transfer test, the four novel colors were incorporated into training trials, and the monkeys were trained with all eight colors with the correction procedure. Training continued until the monkeys performed ≥ 80% correct, at which point the correction procedure was removed. The monkeys then continued training until they once again performed ≥ 80% correct. Then, the monkeys were tested with novel variable delays over the course of 24 sessions. The original training delay was intermixed with novel delays of 100, 200, 400, 800, 1600, 3200, and 6400-ms such that there were twelve trials of each delay per session. Reinforcement was provided following all correct responses.

Color-change detection with novel shapes. Three tests evaluated the monkeys' ability to judge color changes with novel shapes. Each test lasted six sessions, with fourteen test trials per session interleaved with 82 baseline training trials. Example trials from all three tests are depicted in Figure 3.3. In Shape Test 1, one novel shape was displayed in two different colors on a given trial. After the delay, one of the stimuli changed in color. In Shape Test 2, two different shapes were displayed in two different colors and one changed color after the delay. Finally, in Shape Test 3, two different shapes were displayed in two different colors in the sample display, and after the delay, one changed in both shape and color.

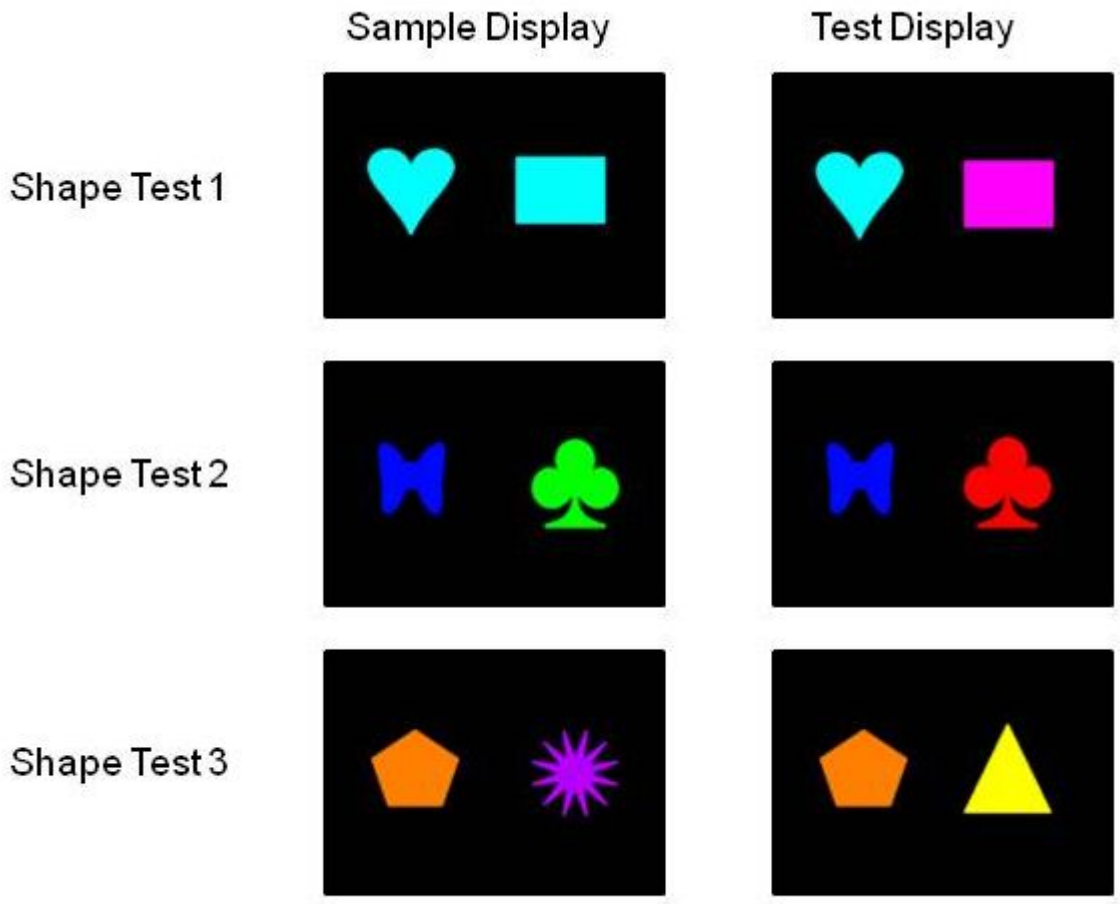


Figure 3.3: Color Changes with Novel Shapes Trial Types (Shape Tests 1-3).

Shape-change detection. In Shape Tests 4, 5A, and 5B, the monkeys were tested to see if they would spontaneously transfer to shape changes after having only been trained with color changes. Example trials from all three tests are depicted in Figure 3.4. Both Tests 4 and 5 were tested for six sessions with fourteen test trials per session. However, in Test 5 there were two trial types (5A and 5B) with seven trials of each type tested per session. Trials in Shape Test 4 included two shapes presented in two different colors, and then following the delay, one changed shape (but not color). In Shape Test 5A, two stimuli of the same shape and color were presented in the sample display, and then after the delay one changed in

shape. Lastly, in Shape Test 5B, two different shapes were displayed in the same color, and after the delay, one changed shape.

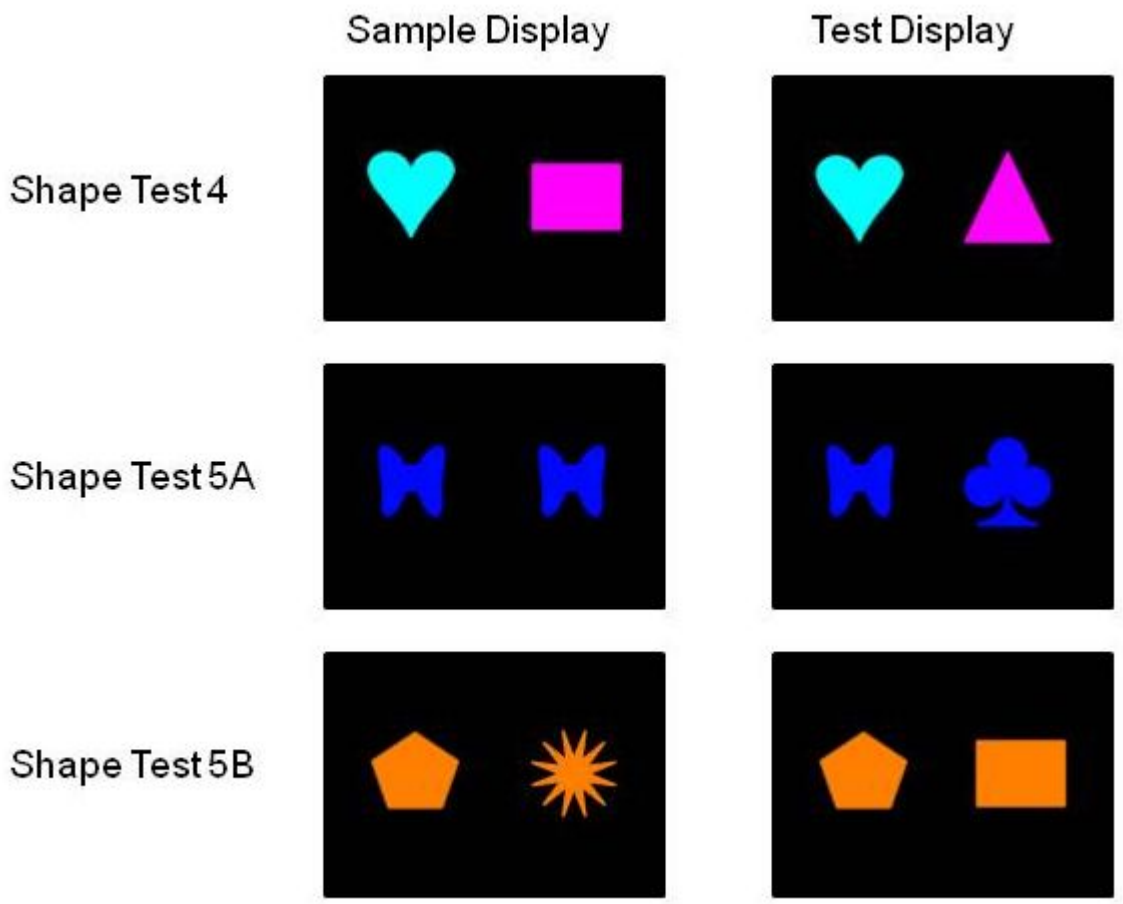


Figure 3.4: Shape Change Trial Types (Shape Tests 4, 5A, 5B).

Location change detection. The monkeys were tested for seven sessions (with twelve test trials per session) with changes in location. In this test, two colored circles were presented in the sample display, and after the delay, one of the circles moved to a new location (Figure 3.5). The monkeys received reinforcement for touching the circle that had changed location.

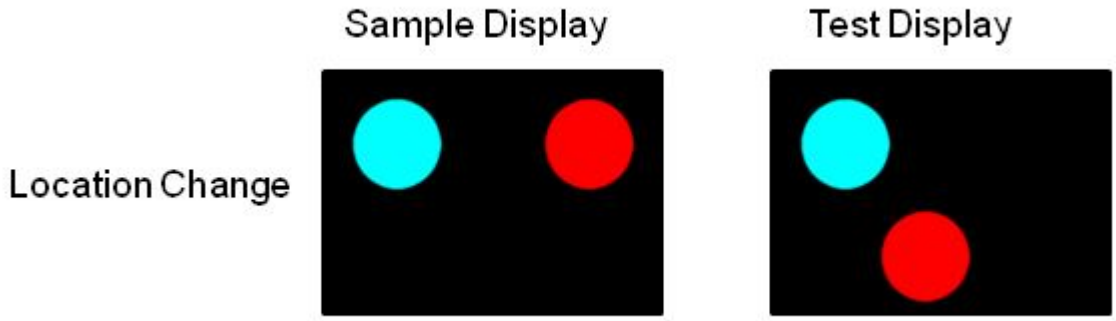


Figure 3.5: Location Change Detection Test Example Trial

Size change detection. The monkeys were next tested with 25% changes in size for six sessions (with fourteen test trials per session). The colored shapes were used in this test and after the delay, one shape either increased or decreased in size by 25% (seven trials of each per session). Example trials are shown in Figure 3.6. The monkeys received reinforcement for touching the shape that had changed size.

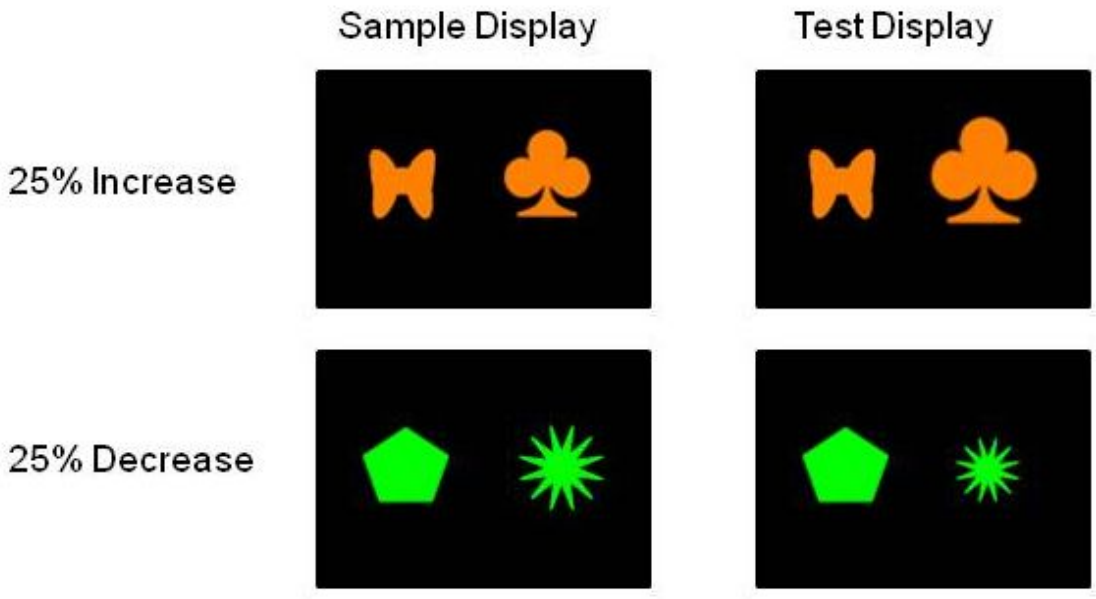


Figure 3.6: Size Change Detection Test Trial Types

Results

Acquisition

M1 and M2 met criterion for acquisition in 53 and 51 sessions respectively. As shown in Figure 3.7, both monkeys showed a rise in performance to 81% correct (early on for M2 and later for M1), but this occurred during their training with a 0-ms delay and was not counted towards the acquisition criterion.

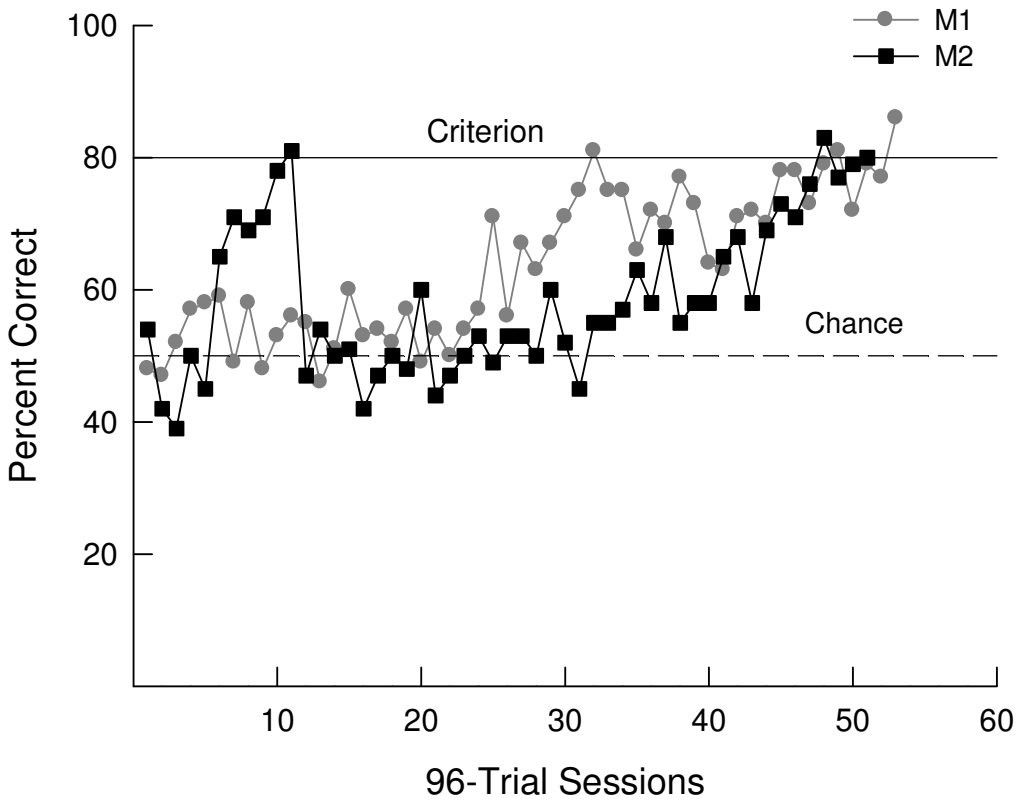


Figure 3.7: Acquisition of the Change Detection Task.

Interestingly, both monkeys frequently made touch responses to the sample display prior to the change (mean of 2.49 ± 0.32 times per trial). An analysis was performed to determine if touch responses to the sample array influenced the monkeys' choices after the stimulus change. Specifically, the analysis questioned whether performance would differ if the

monkey had been touching the object that would change or the object that would remain unchanged. Interestingly, the monkeys were different in this regard. M2 did not perform differently based on the object he had been touching prior to the change (81.7% for changed item vs. 83% for unchanged item; paired sample *t*-test, $t(2) = 0.227$, $p = 0.841$). M1, however, performed better (88.4% correct) when he had been touching the unchanged item vs. the changed item (74.2% correct). This difference was significant (paired samples *t*-test, $t(2) = 7.407$, $p = 0.02$).

Color Transfer

As shown in Figure 3.8, both monkeys showed transfer equivalent to baseline when tested with novel colors as determined by paired samples *t*-tests (M1: $t(5) = 1.09$, $p = 0.33$; M2 $t(5) = 0.48$, $p = 0.65$). M1 and M2 averaged 72.2% and 83.3% correct respectively with novel colors and 78.8 and 80.7% correct respectively on baseline trials. This good transfer with novel colors was not the result of learning across the six test sessions, because first session performance for M1 and M2 was 83.0% and 92.0% correct respectively both of which are significantly greater than chance (binomial tests, $p_s \leq 0.01$).

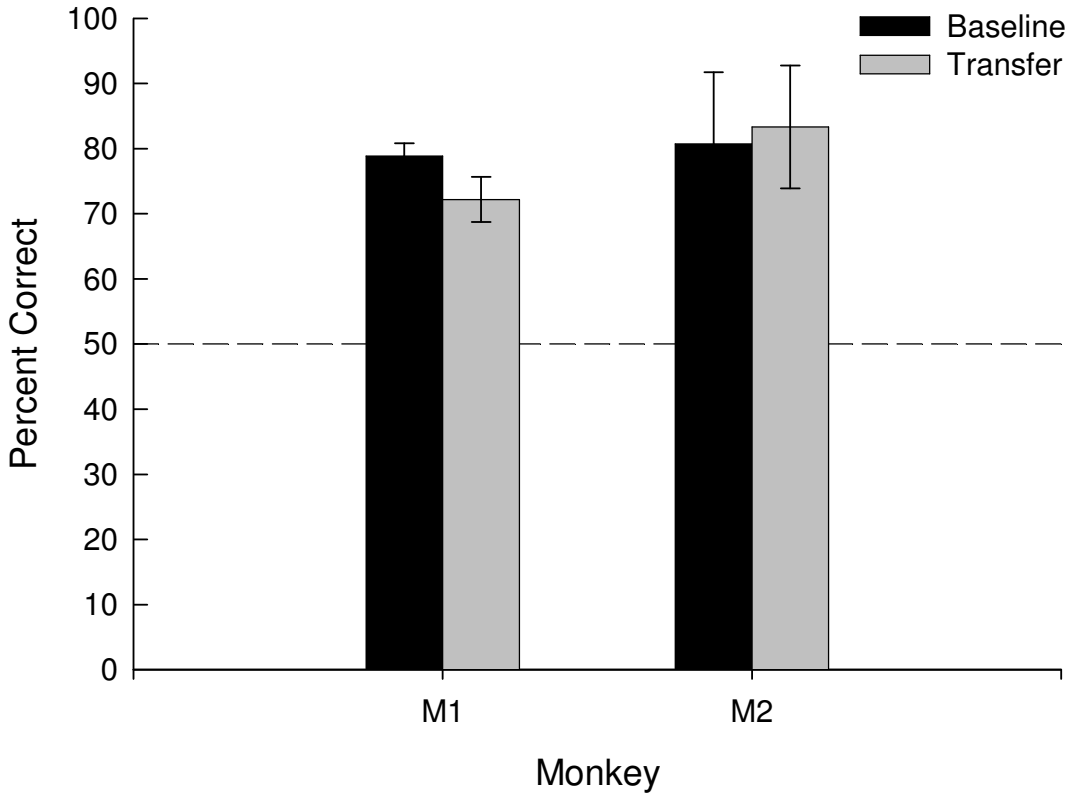


Figure 3.8: Novel Color Transfer Test Performance. Error bars represent standard error of the mean.

Variable Delay Testing

Figure 3.9 shows performance from the variable delay test. Statistical analyses indicated the presence of a significant effect of delay for both monkeys [Separate one-way repeated measures ANOVAs: M1 – $F(7,23) = 8.02$, $p < 0.001$; M2 – $F(7,23) = 9.20$, $p < 0.001$]. In short, performance decreased as a function of delay, as delay increased from 50 to 6400 ms. Correlation analyses demonstrate that learning did not occur across the 24 test sessions, as mean performance was not significantly correlated with session for either monkey (M1: $r = -0.31$, $p = 0.14$; M2: $r = 0.18$, $p = 0.39$). Single sample *t*-tests against chance (50%) were conducted to determine if performance was significantly greater than chance at all delays tested, which indeed they were (all $t_s(23) \geq 2.46$, all $p_s \leq 0.02$). To summarize, the analysis

demonstrates that the monkeys' performance was stable across sessions and significantly greater than chance at all delays. This indicates that training was not necessary for the monkeys to perform change detection at delays longer than their 50-ms training delay.

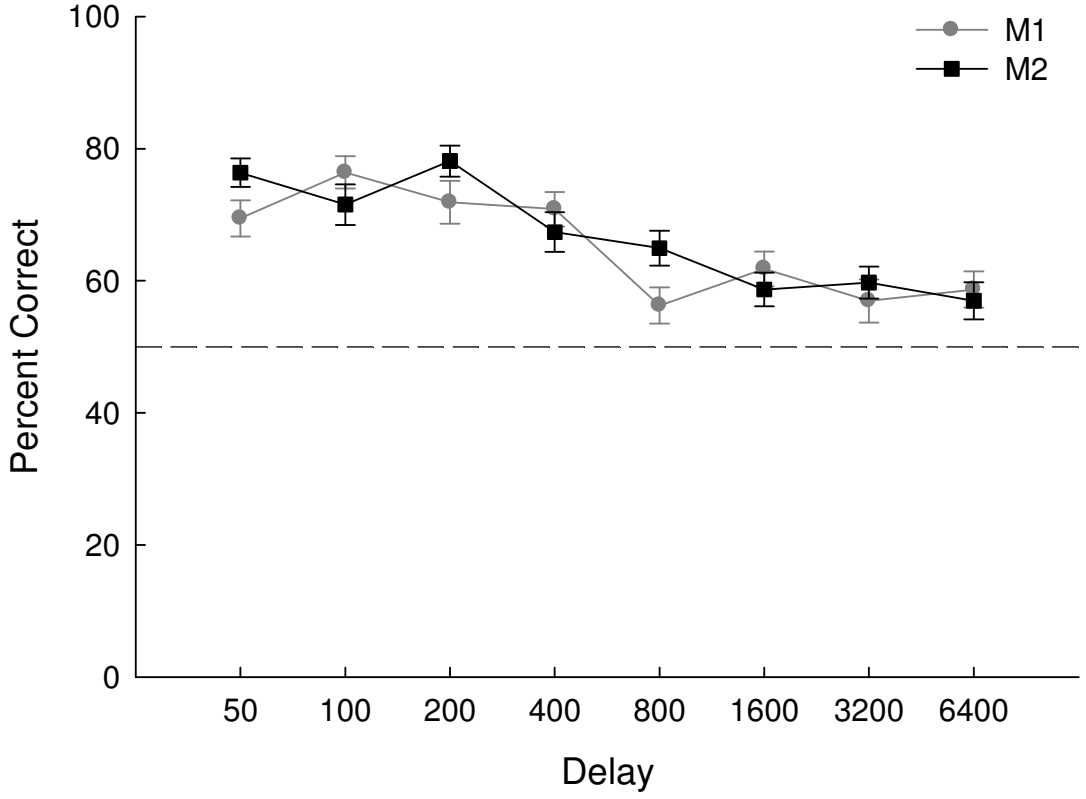


Figure 3.9: Variable Probe Delay Test Performance. Error bars represent standard error of the mean.

Color-change detection with novel shapes. The results of the three color-change detection with novel shapes tests (Shape Tests 1-3) are shown in Figure 3.10. Both monkeys performed significantly better on baseline trials than on transfer trials in Shape Test 1 [paired sample *t*-tests; M1: $t(5) = 3.92$, $p = 0.01$; M2: $t(5) = 2.49$, $p = 0.05$]. In fact, neither monkeys' transfer performance was statistically better than chance (50%) [single sample *t*-tests: M1: $t(5) = 0.56$, $p = 0.6$; M2: $t(5) = 0.34$, $p = 0.75$]. In Shape Test 2, M1's transfer performance was not significantly different from baseline [$t(5) = 0.68$, $p = 052$]. M2, however, performed significantly

better on baseline trials than on transfer trials [paired sample *t*-test, $t(5) = 2.82$, $p = 0.04$]. M2's transfer performance in Shape Test 2 (61.9%) was significantly better than chance [$t(5) = 2.57$, $p = 0.05$]. Neither monkey showed a significant difference in performance between baseline and transfer in Shape Test 3 [M1: $t(5) = 1.08$, $p = 0.33$; M2: $t(5) = 1.60$, $p = 0.17$].

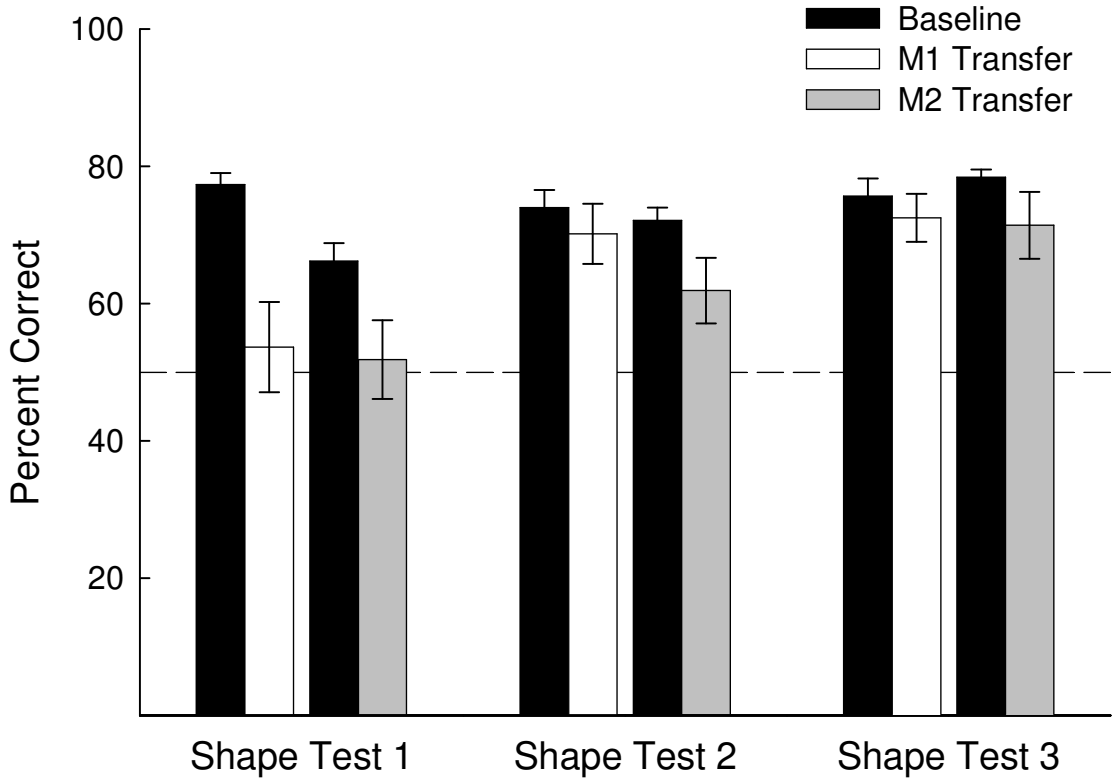


Figure 3.10: Color Change with Novel Shapes Performance. Error bars represent standard error of the mean.

Shape-change detection. The results of the shape-change detection tests (Shape Tests 4 & 5) are displayed in Figure 3.11. In Shape Test 4, both monkeys showed performance equivalent to baseline (full transfer) on transfer trials [M1: $t(5) = 0.89$, $p = 0.41$; M2: $t(5) = 0.49$, p 0.64]. The monkeys again did not show significant differences between baseline and transfer trials in Shape Test 5A [M1: $t(5) = 2.29$, $p = 0.07$; M2: $t(5) = 0.44$, $p =$

0.68]. However, both monkeys performed significantly worse on transfer trials than on baseline in Shape Test 5B [M1: $t(5) = 3.74$, $p = 0.01$; M2: $t(5) = 6.45$, $p = 0.001$]. M1's performance (54.76%) was not significantly different from chance [$t(5) = 1.10$, $p = 0.32$]. M2's performance (61.87%) was significantly better than chance [$t(5) = 4.03$, $p = 0.01$].

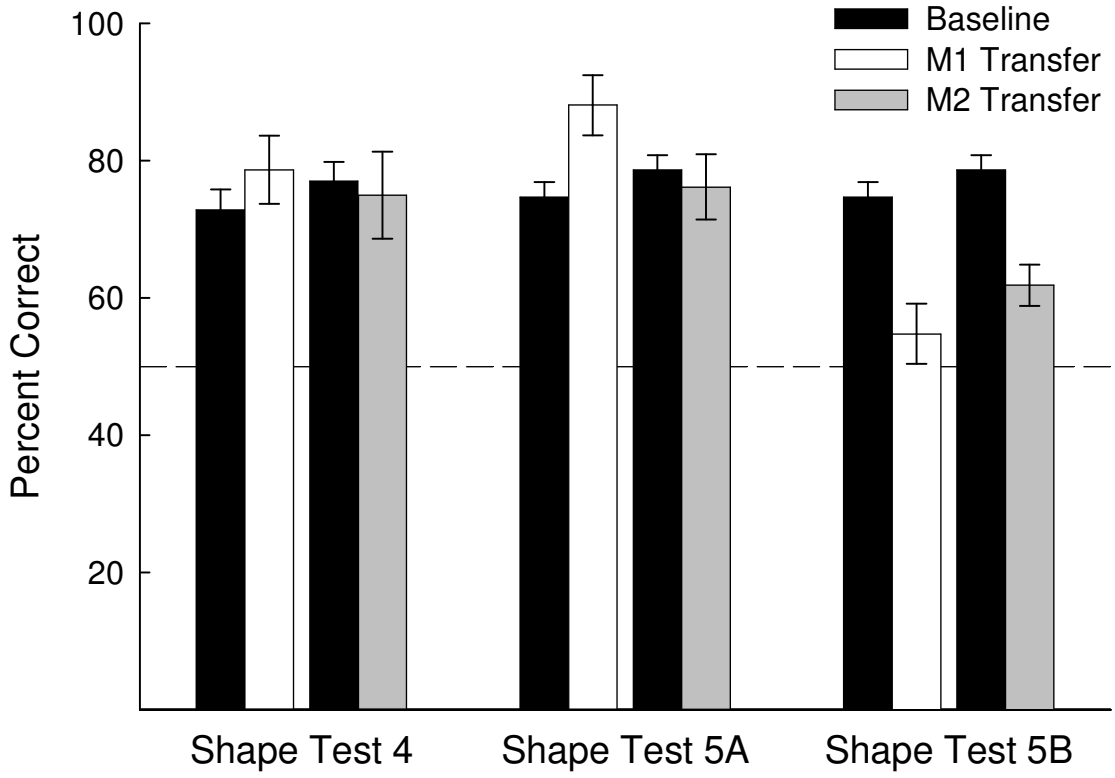


Figure 3.11: Shape Change Test Performance. Error bars represent standard error of the mean.

Location-change detection. As shown in Figure 3.12, in the Location Change Test, both monkeys' transfer performance was not significantly different from baseline [M1: $t(6) = 1.28$, $p = 0.24$; M2: $t(6) = 0.65$, $p = 0.54$].

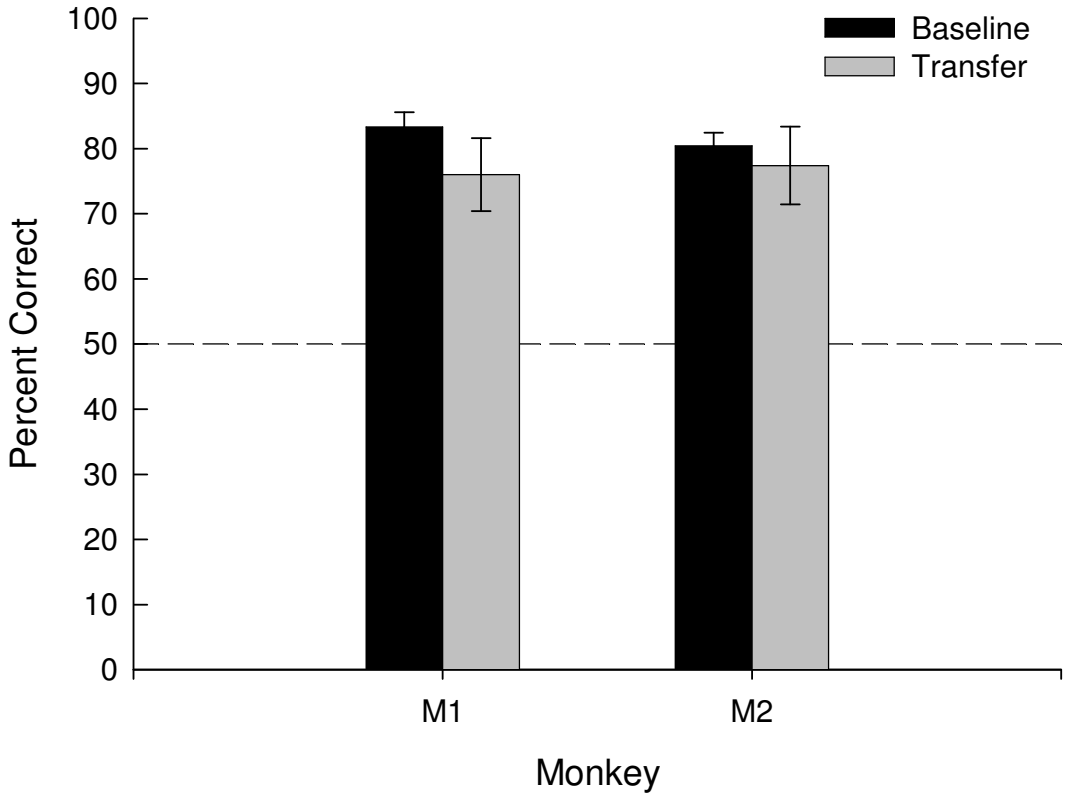


Figure 3.12: Location Change Test Performance. Error bars represent standard error of the mean.

Size-change detection. As shown in Figure 3.13, both monkeys performed significantly worse on size change trials than on baseline trials [M1: $t(5) = 5.89$, $p = 0.002$; M2: $t(5) = 5.34$, $p = 0.003$. In fact, M1's size change performance (55.95%) was not significantly different from chance (50%) [single sample *t*-test, $t(5) = 1.05$, $p = 0.34$]. However, M2's size change performance (61.90%) was significantly better than chance [$t(5) = 4.03$, $p = 0.01$].

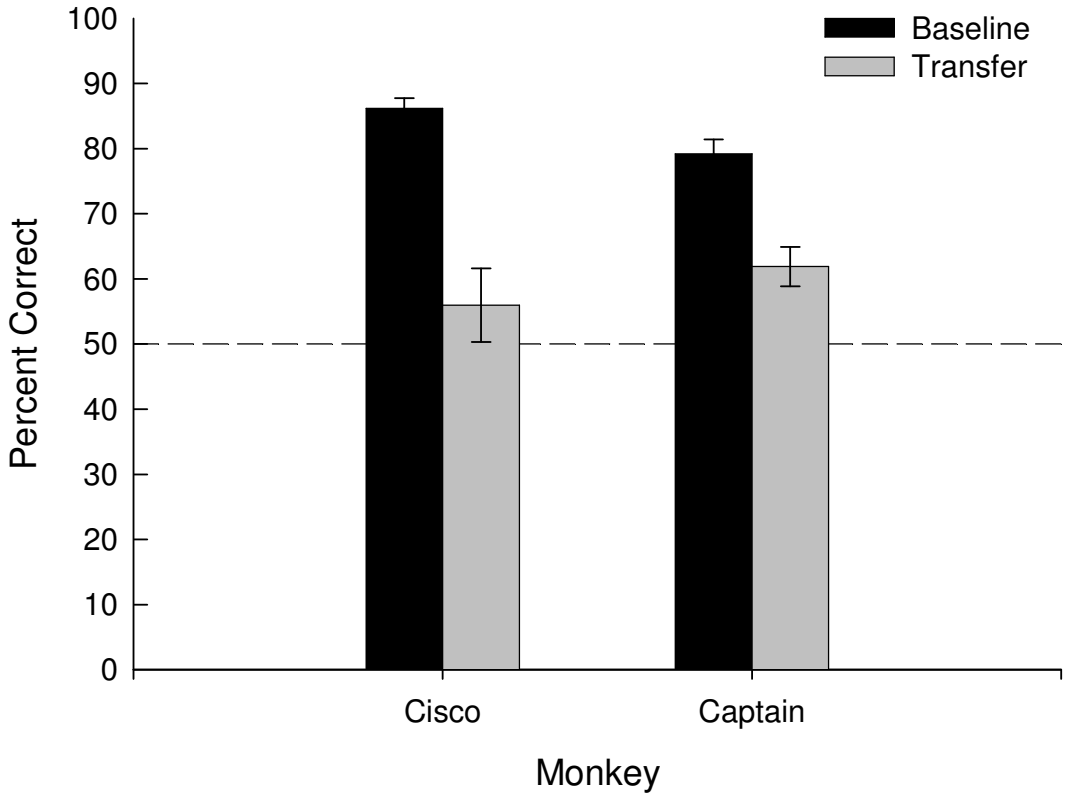


Figure 3.13: Size Change Test Performance. Error bars represent standard error of the mean.

Discussion

Both monkeys learned to perform the task and met criterion for acquisition in a similar number of sessions (51 and 53 sessions). This time to acquisition is comparable to what has been shown in other monkey memory tasks like same/different and delayed matching to sample (e.g., Katz et al., 2002; Mishkin & Delacour, 1975; Wright et al., 2003; Wright, 1999) . Both monkeys also showed full transfer (transfer equivalent to baseline) to novel color stimuli and mostly good transfer to color changes with novel shapes, demonstrating that learning was not tied to the four different colored training circles. However, it should be noted that the monkeys performed at chance (52.8%) in Shape Test 1, which was their first experience with novel shapes. The monkeys may have been averse to the novelty of the shapes, or confused

by the sample displays which included two stimuli of the same color (the monkeys had never seen a sample display with both stimuli presented in the same color before this). However, they quickly overcame this aversion as they both performed very well in Shape Tests 2 and 3 which also tested them for their ability to detect color changes with novel shapes.

Importantly, when the monkeys were abruptly tested with delays longer than their 50-ms training delay (ranging from 100 to 6400-ms), they performed above chance at all delays. The 50-ms delay falls within the time frame of attentional capture as studied in humans. However, the monkeys' performance did not fall abruptly as delays increased beyond 800-ms, which are delays longer than the limits of attentional capture (Cusack et al., 2009; Pashler, 1988). In fact, their performance gradually declined as the delay interval increased, a result that would be expected in a memory task. Thus, it is highly unlikely that the monkeys' performance of the change detection task was mediated by an attentional capture mechanism, but rather must have been mediated by VSTM.

In addition, tests with shape changes and location changes showed that the monkeys could readily transfer to novel types of change. This was a particularly important result because it demonstrated that despite the fact that the monkeys were only explicitly trained to detect color changes, they were able to spontaneously transfer their knowledge of the task to the novel changes of shape and location. Transfer was equivalent to baseline in Shape Tests 4 and 5A and in the location change test. It should be noted, however, that both monkeys did not perform very well in Shape Test 5B. It is unclear why the monkeys had trouble in this test, as they had already demonstrated good performance with shape changes in Shape Tests 4 and 5A.

The monkeys did not transfer well in the size change test, their performance was at (M1) or near (M2) chance. The size change test was atypical compared to the others tests because it was the only test in which identical (in shape and color) stimuli were presented in the test display in identical locations as the sample display. Furthermore, the difficulty of this

test may have been confounded by the somewhat subtle (25% increase/decrease) size changes used. In a follow-up test with M1 using 50% size changes, his performance was 73.57% which was not significantly different from his baseline performance of 81.67% [$t(5) = 1.07$, $p = 0.33$]. Thus, it seems that through some combination of the additional experience and more salient size changes, at least one monkey was able to accurately transfer to size changes.

To summarize, this experiment demonstrated that monkeys readily learned to perform the change detection task, performed it using short-term memory, and developed some generalized concept of "change" as evidenced by their good performance with novel changes. These findings confirm that the change detection task originally developed for use with humans is also a suitable memory task for rhesus monkeys. Although the monkeys had to learn the rules of the task through the contingencies of reinforcement, they nonetheless demonstrate the ability to perform the task in an analogous way to humans. In addition, their ability to perform with multiple types of change is advantageous, as most human change detection tasks employ multiple stimulus/change types like colored squares, random polygons, and Snodgrass drawings (e.g. Alvarez & Cavanagh, 2004; Eng, Chen, & Jiang, 2005). Thus, a direct comparison of human and monkey change detection performance is possible, using similar parameters to those tested previously in the literature.

CHAPTER 4: VISUAL SHORT-TERM MEMORY IN RHESUS MONKEYS

Introduction

Change detection is a task that is imminently suitable for studies of animal memory. We demonstrated in Chapter 3 that rhesus monkeys readily learned the change detection task and solved the task by looking for change in a general sense, indicating that they perform the basic change detection task in a manner analogous to a human subject. Change detection has been shown to test non-verbal, visual memory such that a lack of verbal processing by rhesus monkeys should not confer a disadvantage relative to humans (Luck & Vogel, 1997; Alvarez & Cavanagh, 2004). Also, two other groups have tested rhesus monkeys with change detection tasks and found good performance with large memory displays (Heyselaar et al., 2011; Buschman et al., 2011). However, neither of these groups compared their monkey results to human change detection performance.

Testing rhesus monkeys in change detection also provides the opportunity to evaluate the two competing models of VSTM, the fixed-capacity model and the continuous-resource model for the first time in a non-human species. A finding of converging evidence from both monkeys and humans favorable to one model might improve our understanding of VSTM in general. Improving our conceptual understanding of VSTM could guide future investigations that seek to understand the neural basis of VSTM, and many of these studies could be carried out in rhesus monkeys. They are the premier medical model for humans, and invasive studies such as lesions, electrophysiological recordings, inactivation, stimulation, and pharmacological and neurotransmitter manipulations could be performed on rhesus monkeys.

Studies with rhesus monkeys performing visual list memory tasks have shown qualitative similarities between monkeys and humans. Both species show serial position functions with primacy and recency effects that depend on the delay (Wright et al., 1985; Wright, 2007). Possibly, rhesus monkeys would show qualitative similarities to human in change detection as well. Thus, the purpose of this experiment was to test rhesus monkeys in the change detection task with similar parameters to those tested with humans in Chapter 2

(Experiment 1). The same rhesus monkeys (Chapter 3) were tested with the same stimuli and some of the same display sizes used to test human subjects in Chapter 2.

Methods

Subjects

The subjects were the two adult male rhesus monkeys (described in Chapter 3). All animal procedures were in compliance with the National Institute of Health guidelines and were approved by the University of Texas Health Science Center at Houston Institutional Animal Care and Use Committee.

Apparatus

The apparatus was the same as described in Chapter 3.

Stimuli

The stimuli were eight 4-cm diameter colored circles (aqua, blue, green, magenta, orange, purple, red, yellow) and 976 different clip art images. Example clip art images are shown in Chapter 2 and the colored circles are displayed in Chapter 3. The stimuli were displayed on the same 4×4 grid described in Chapter 3 and subtended a visual angle of approximately 5.75 degrees based on the average distance of the monkey from the screen.

Training and Test Procedures

Colored circles. Following the completion of the training and testing sessions described in Chapter 3, the monkeys were trained for 65 sessions (M1) and 18 sessions (M2) with increasingly larger display sizes. They first trained with sessions containing three items per displays (57 sessions for M1, 10 sessions for M2). M1 was trained with many more sessions because he had started the experiment earlier than M2 and was training on 3-item displays during the period in which the software for testing intermixed display sizes was developed. Performance was comparable between the two monkeys despite the difference in training duration as M1's final day of training performance was 82% correct and M2's performance was

77%. They were next trained with four sessions of intermixed trials with display sizes of 2, 3, and 4 and then four sessions of intermixed trials with display sizes of 2, 3, 4, and 5. Finally, they were tested for four sessions with display sizes of 2, 3, 4, 5, and 6. The results from the display sizes of 2, 4, and 6 are analyzed here. Across the four sessions, the monkeys were tested with a total of 76 trials per display size 2, 4, and 6.

Clip art. After the completion of the colored circles tests, the monkeys were gradually transitioned to performing the task with clip art images. They were first tested for six sessions with twelve clip art transfer trials inserted, as in the transfer tests of Chapter 3. Both monkeys transferred well; baseline performance for M1 and M2 was 85.3% and 85.2% correct respectively and transfer performance was 74.8% and 69.4% correct respectively. However transfer performance was not statistically equivalent to baseline [M1: $t(5) = 3.57$, $p = 0.02$; M2: $t(5) = 3.20$, $p = 0.02$], such that additional training with clip art images was necessary. Following the transfer test, the monkeys were gradually transitioned to performing the task with all clip art images over the course of three sessions. Across these three 96-trial sessions, the number of clip art trials was increased from 32 to 48 to 64 and the number of colored circle trials was decreased accordingly. On the fourth session, the monkeys began performing the two-item display change detection task entirely with clip art images. After 11 sessions (M1) and 10 sessions (M2) the display sizes were again gradually increased over the course of three sessions. During the course of these training sessions both monkeys achieved a performance criterion of 80% correct or greater. The display size increase occurred more quickly than with colored circles because the monkeys were accustomed to large display sizes at this point. On the fourth session, testing with display sizes of 2, 4, and 6 clip art objects began.

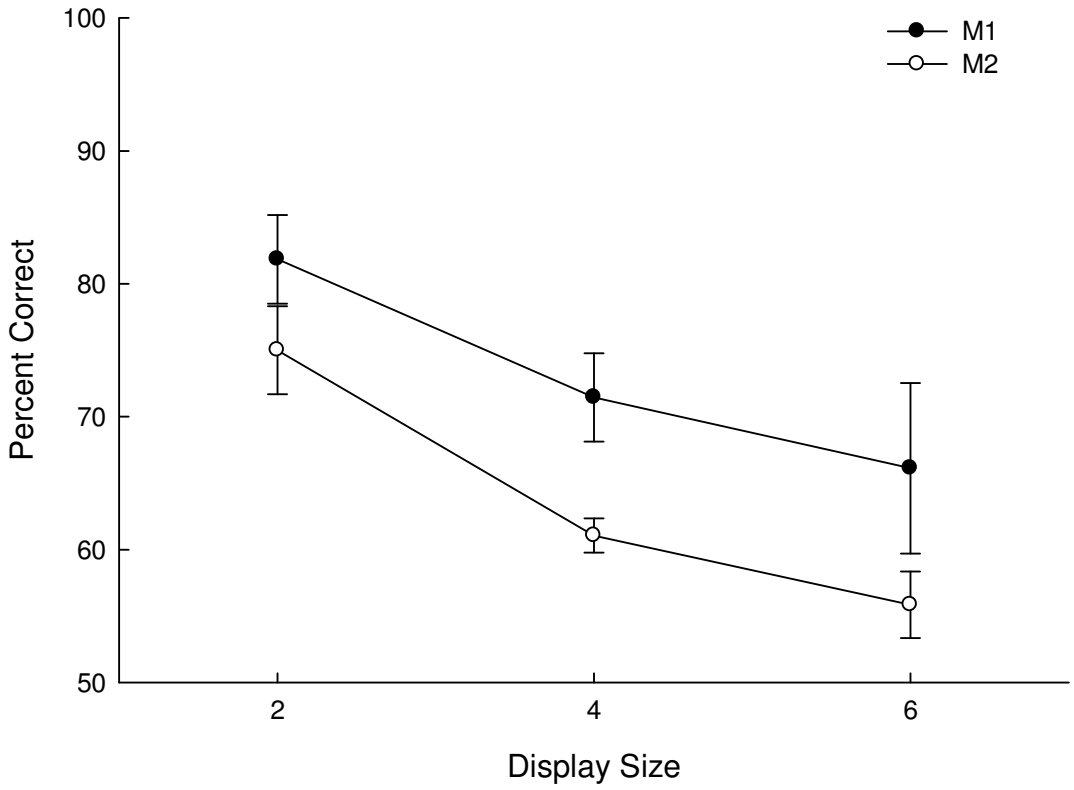
The different display sizes were randomly intermixed. The images were selected without replacement such that they were trial unique for two sessions. A total of eighteen 96-trial sessions were tested. A total of 144 trials per display size were included in the analysis.

In these trials the monkeys had a viewing time of 5000 ms and a 50-ms delay as in their training in Chapter 3. Other trials with viewing times ranging from 1000 to 4500 ms and delays ranging from 200 to 1000ms were tested but were not included in the analysis of this experiment in order to make it comparable to the test conducted with colored circles.

Results

As predicted, the monkeys' performance was high with two-item displays but fell as the display size increased for both colored circles and clip Art (Figure 4.1). Separate repeated-measures ANOVA of display size × stimulus type showed a significant effect of display size for both monkeys [M1: $F(2,6) = 20.258$, $p = 0.002$; M2: $F(2,6) = 12.469$, $p = 0.007$). In addition, M2 showed a significant effect of stimulus type [$F(1,3) = 11.14$, $p = 0.04$], but M1 did not. M2 performed better with clip art stimuli than with colored circles.

A



B

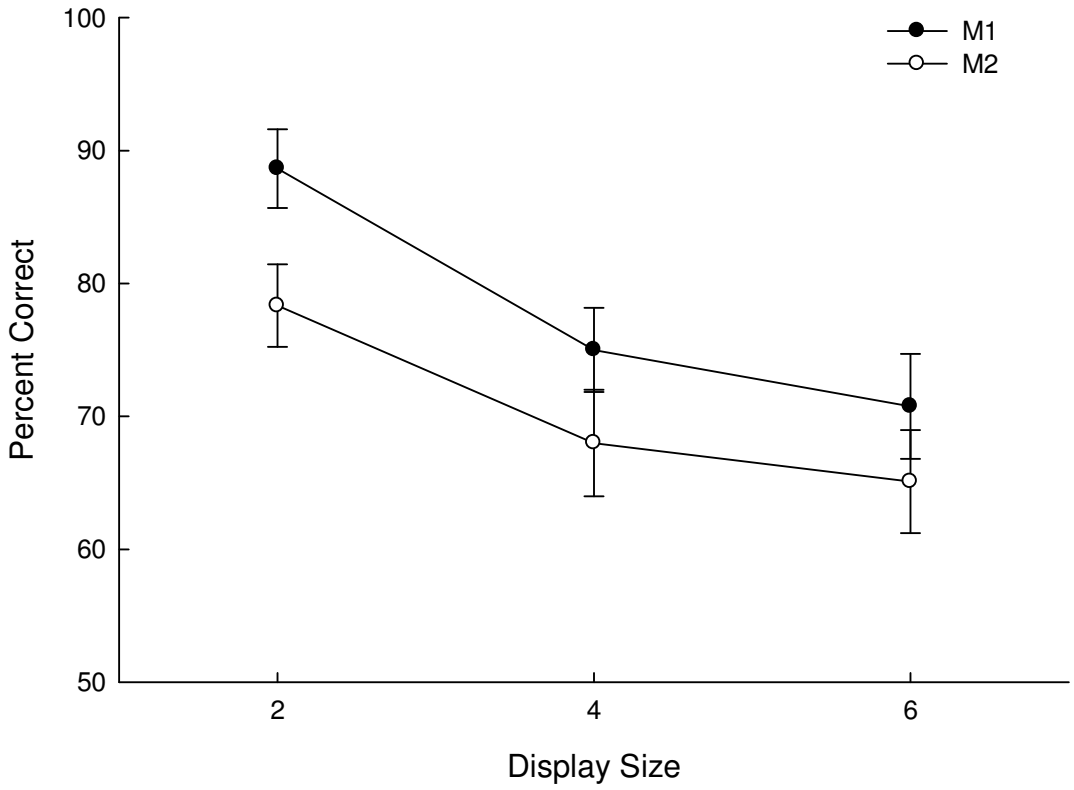


Figure 4.1. Percent Correct in Change Detection Task for A) Colored Circles and B) Clip Art. Error Bars Represent Standard Error of the Mean.

Estimating capacity. Capacity measures were calculated using Equation 2.1 (Chapter 2). Mean capacity estimates for each stimulus type and display size are displayed in Figure 4.2. Mean capacity for colors was found to be 0.71 ± 0.24 and mean capacity for clip art was 1.02 ± 0.19. Thus the monkeys' VSTM capacity was found to be approximately one item or less.

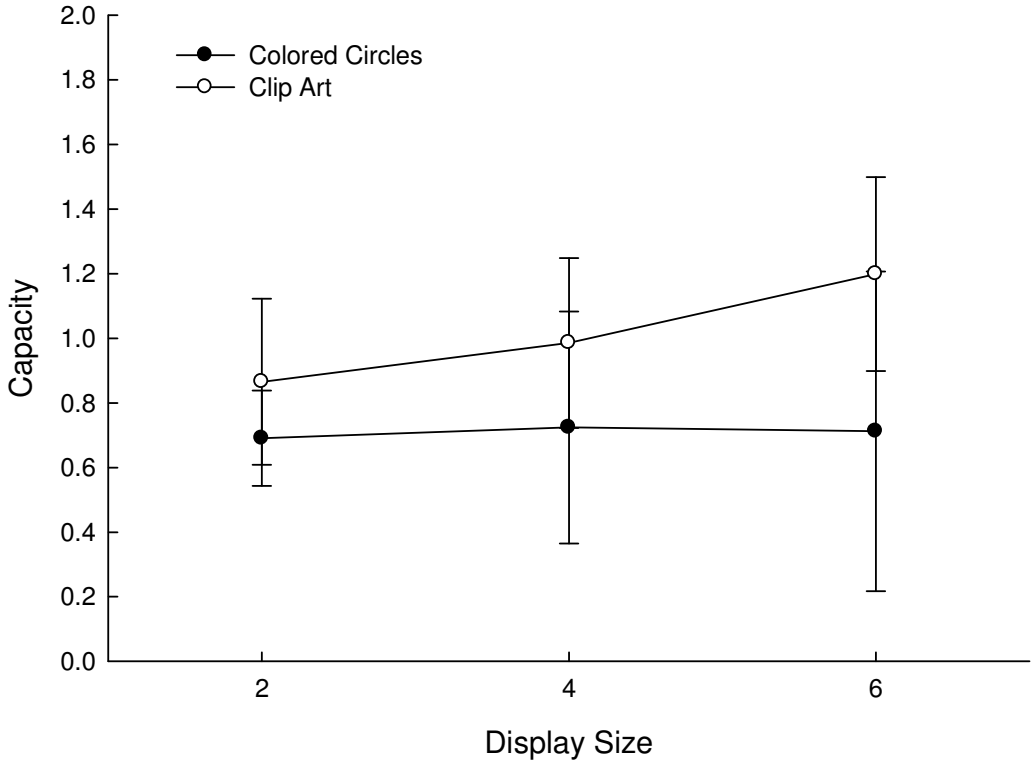


Figure 4.2: VSTM Capacity Estimates for Rhesus Monkeys with Colored Circles & Clip Art.

Color analysis. Ten 96-trial sessions of two-item display change detection were analyzed to determine the extent to which the monkeys confused similar colors. These data were collected before the tests with two, four, and six item displays described above. This analysis sought to determine whether or not the monkeys were more likely to make mistakes when one color changed to a similar color vs. when one color changed to a less similar color. In order to test this, a multidimensional scaling analysis was performed. This algorithm works by transforming the 8-dimensional accuracy matrix (8 colors) into a 2-dimensional space. This 2-dimensional perceptual space displays the colors on two axes which maintain the distance structure in the original matrix as well as possible. Thus, in reading the plot (Figure 4.3) the arbitrary units of "distance" are tied to performance when one color changed to another.

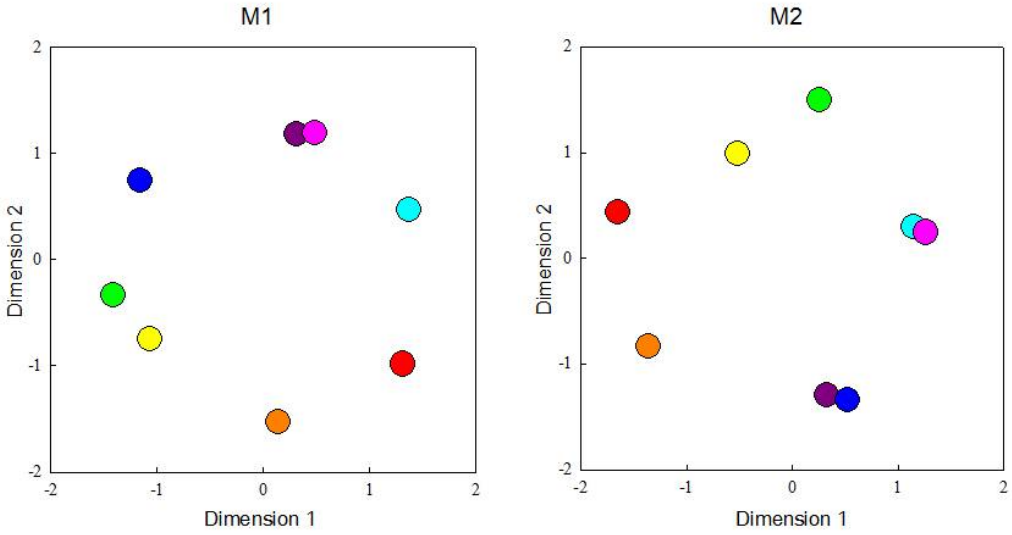


Figure 4.3. Multidimensional Scaling of Performance for Sample Colors Changing to Test Colors.

Small distances between colors indicate that performance was low when one color changed to another. For instance, M1 performed at chance when magenta changed to purple and vice versa (52% correct). Likewise, M2's performance was near chance when blue changed to purple and vice versa (57%). However, M1's performance was perfect when red changed to green and vice versa (100% correct) and M2's performance was perfect when green changed to orange and vice versa (100% correct). A large proportion of the variance is accounted for by color confusion as r^2 values were 0.61 for M1 and 0.56 for M2. Stress values were 0.303 and 0.332 for M1 and M2 respectively.

Continuous-Resource Model. As with humans, we computed *d'* values from the monkeys' performance using Equation 2.2. Mean *d'* values for both stimulus types are plotted in Figure 4.4.

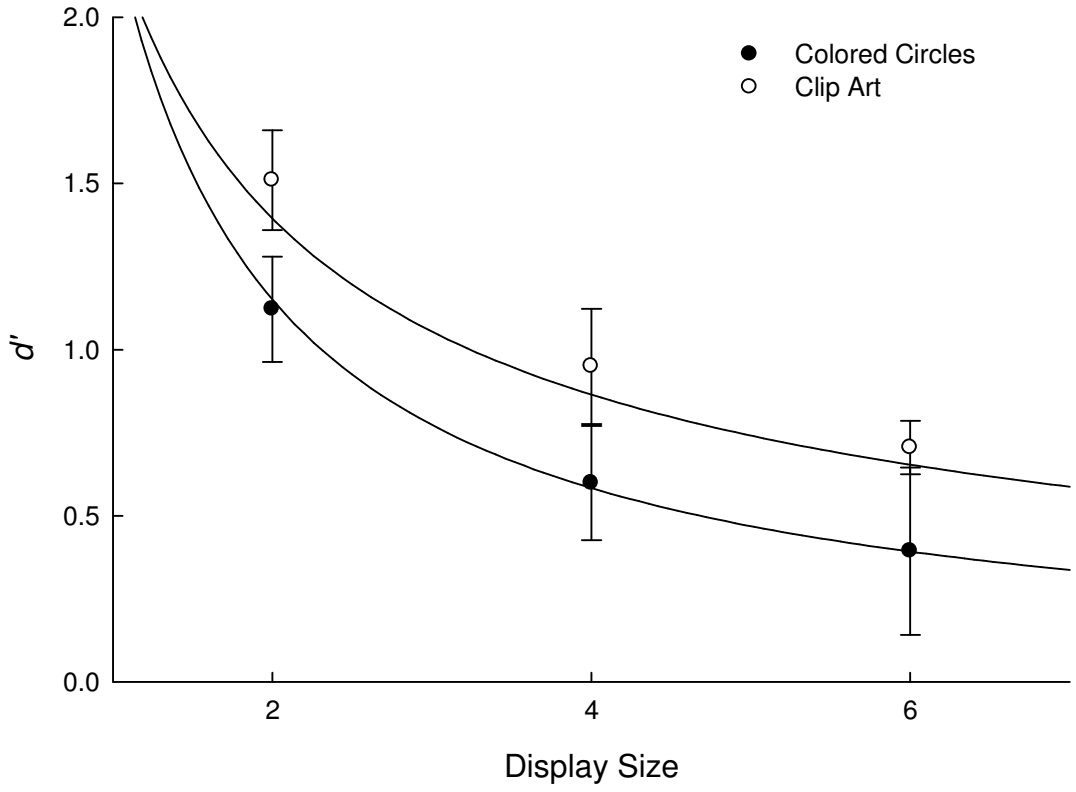


Figure 4.4: Power Law fits for *d'* for Colored Circles and Clip Art.

Power law functions were generated as described in Chapter 2. For both stimulus types the power law functions provided a good fit to the *d'* values; r^2 values were 0.98 and 0.99 for colored circles and clip art respectively. These r^2 values were extremely significant [Colored Circles: $F(1,4) = 165.76$, $p = 0.0002$; Clip Art: $F(1,4) = 11143.51$, $p < 0.0001$].

Comparison to Human Subjects. The data from rhesus monkeys were compared to the data collected from human subjects in Chapter 2, Experiment 1. Overall performance is compared in Figure 4.5. On average, humans outperformed monkeys by 16.5% on clip art

trials and by 22.0% on colored circle trials. A repeated-measures ANOVA of display size × stimulus type × species was conducted. This ANOVA demonstrated that there was a significant effect of display size [$F(2,24) = 39.045$, $p < 0.001$], a significant effect of species [$F(1,12) = 60.159$, $p = 0.001$], as well as a significant interaction of stimulus type and species [$F(1,12) = 6.679$; $p = 0.024$]. The interaction results from the fact that one monkey (M2) performed significantly better with clip art than with colors.

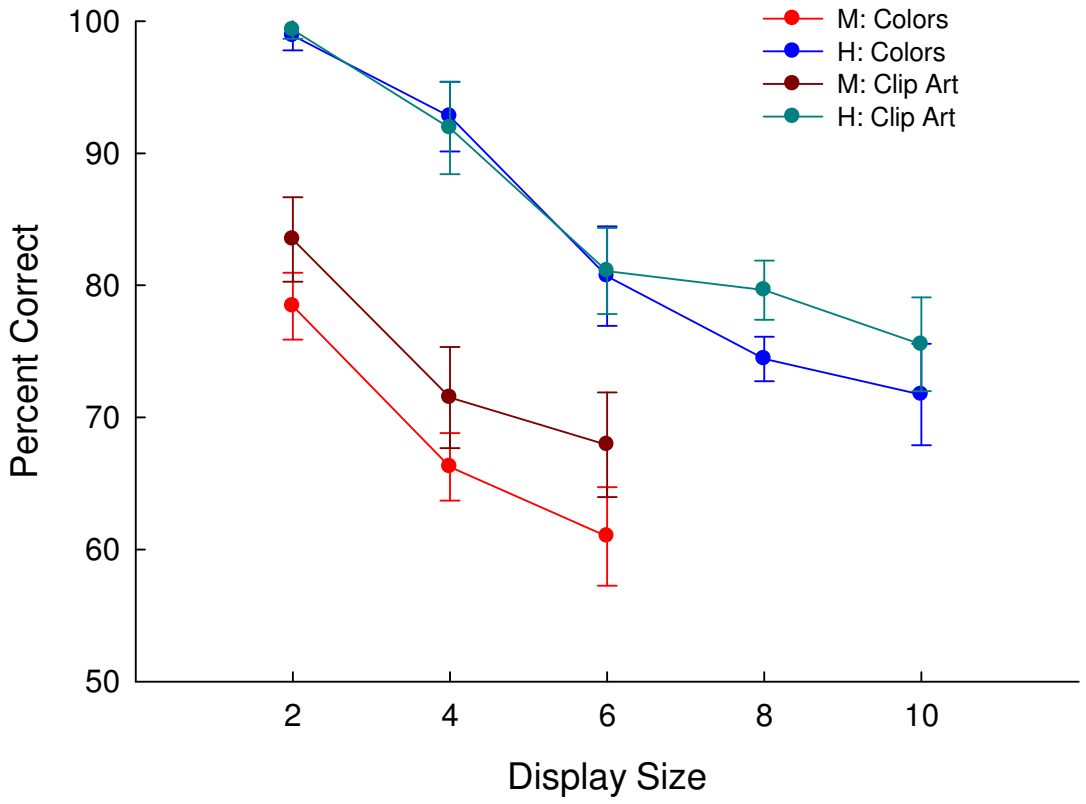


Figure 4.5: Comparison of Change Detection Performance by Humans and Rhesus Monkeys with Colors and Clip Art.

Capacity Estimates. Capacity estimates between the two species were also compared (Figure 4.6). Mean capacity estimates for humans were 2.46 ± 0.35 for colors and 2.78 ± 0.39 for clip art, whereas mean capacity estimates for monkeys were 0.71 ± 0.24 and 1.02 ± 0.19

for colors and clip art respectively. Thus, based on a fixed-capacity model of VSTM, humans could store approximately 1.5 more visual items than could rhesus monkeys.

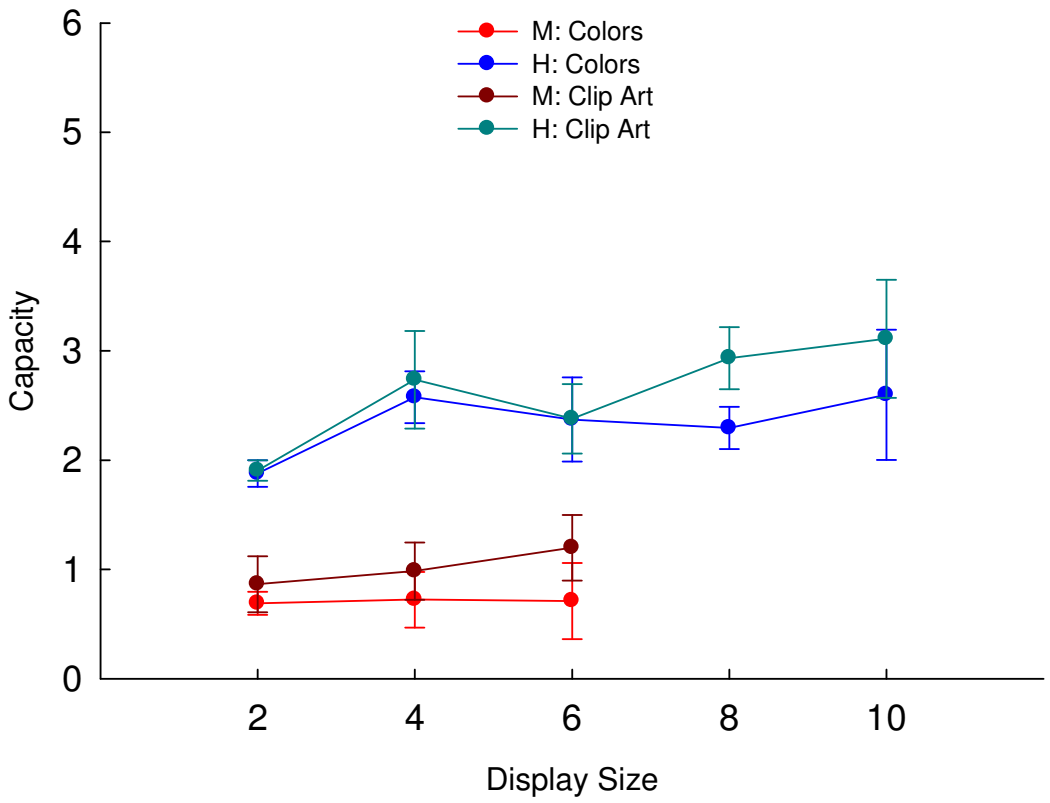


Figure 4.6: Comparison of Capacity Estimates for Humans and Rhesus Monkeys. Error bars represent standard error of the mean.

Continuous-Resource Model. Monkey and human performance was also compared using *d'* measures as per the continuous-resource model. *d'* values and power law fits for both species are compared in Figure 4.7. Although the humans clearly outperformed the monkeys, the exponents of the power law functions fall within a similar range, and the functions actually differ primarily by the coefficient suggesting a similar shape to the functions but difference in absolute level. For clip art, the human exponent was -0.72 and the monkey exponent was -0.69. For colors, the human exponent was -0.86 and the monkey exponent was -0.98. The full equations of the power law functions are listed in the legend of Figure 4.7.

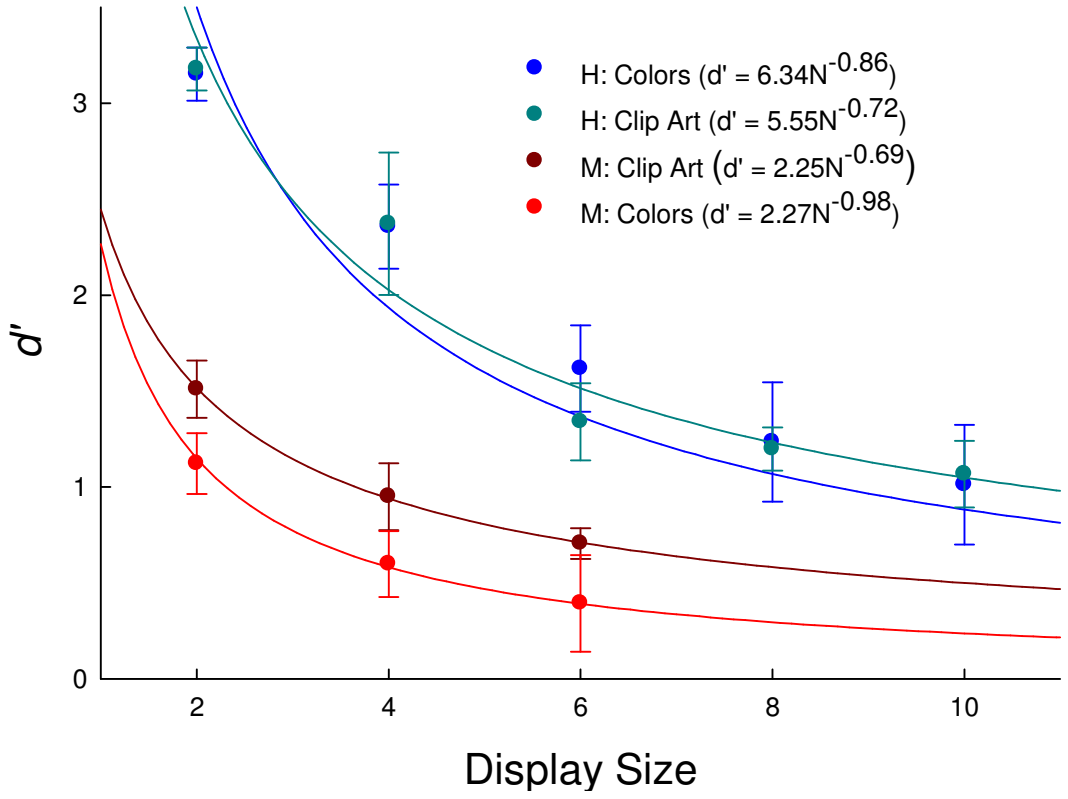


Figure 4.7. Comparison of *d'* values and power law fits for Humans and Rhesus Monkeys.

Discussion

The monkeys performed well in the tests with colored circles and clip art, but the capacity estimates generated from their performance are shockingly low. The mean capacity values of 0.71 ± 0.24 for colors and 1.02 ± 0.19 for clip art indicate that according to a fixed-capacity model of VSTM monkeys can only maintain one item of visual information in VSTM at a time. While it is perhaps not surprising that their capacity limits are lower than those obtained for humans, a limit of a single stimulus seems unusually low. In fact, such a finding is difficult to reconcile with previous work with rhesus monkeys demonstrating that they can

accurately maintain four or more visual or auditory stimuli in memory during a list memory task (Wright, 2007). Although stimuli are presented sequentially in list memory tasks, list memory would still be impossible to perform at that level of accuracy with a VSTM capacity less than or equal to one. In fact, one would predict that performance should be at chance in list memory if capacity were only one.

Another problem with a fixed-capacity model interpretation of VSTM arises from the multi-dimensional scaling analysis done on the colored circle data from monkeys. Fixed-capacity models describe VSTM as a high-resolution storage system in which stimuli are stored perfectly, or not stored at all. The fact that monkeys confused similar colors (e.g. purple and magenta) is not consistent with such a high-resolution storage system because stimuli that are stored perfectly should not be confusable.

The continuous-resource model perhaps provides a more satisfactory framework for VSTM in both rhesus monkeys and humans. As shown in Figure 4.7, both species *d'* values were extremely well fit by power law functions, as predicted by the model. Furthermore, the continuous-resource model provides a good explanation for the color confusion results from monkeys. Because the continuous-resource model predicts noisy representations in memory, it is easy to imagine how a noisy representation of magenta could be confused for purple and likewise for other similar colors.

Another advantage of the continuous-resource model is that it provides a lens through which to see the striking qualitative similarities that occur between monkeys and humans. While the behavioral performance shows the qualitative (and obvious) similarity of a decline in performance as display size increases, the continuous-resource model power law fits allow us to see that the decline in performance can be fit by the exact same type of function with very similar power values for both species.

Although the continuous-resource model allows us to see the qualitative similarities between the two species, there is a difference in time course that should be noted. The

human subjects were tested with 1000-ms viewing times and 900/1000-ms delays. However, the rhesus monkeys were tested with 5000-ms viewing times and 50-ms delays. Thus, a more stringent test of the models and of the qualitative similarities we found between species would be to test the monkeys with the same viewing time and delay used with humans, which will be pursued in Chapter 5. Another minor procedural difference was that humans were tested with a set of six colored squares and monkeys were tested with a set of eight colored circles. This difference will also be addressed in Chapter 5.

CHAPTER 5: CLOSER MATCHED TESTING CONDITIONS FOR RHESUS MONKEYS

Introduction

Chapter 4 demonstrated qualitative similarities in VSTM performance between monkeys and humans, with the same stimuli and overlap in display sizes. Nevertheless, the viewing time and delay period differed between the two species (1000-ms viewing time and delay for humans, 5000-ms viewing time and 50-ms delay for monkeys) and monkeys were tested with a set of eight colored circles, while the humans were tested with a set of six colored squares. The critical difference here is the set size, or the number of items in the stimulus pool used to compose trials. When the monkeys were tested with the eight item set of colored circles in Chapter 4, they were tested with display sizes of two, four, and six. As such, each of the eight colors appeared a maximum of one time per trial. However, since the human trials were drawn from a set of six colors, for display sizes of six, eight, and ten there was at least one color that appeared more than once per trial. The color repetition within trials may make the task more difficult because subjects have to keep track of locations in addition to colors to accurately identify the changed stimulus.

Also, for both the six and eight item color sets the colors repeat across trials. This leads to the development of proactive interference across trials. For instance, Makovski & Jiang (2008) showed using change detection with humans that the repetition of colors across trials results in diminished performance. The same type of effect has been shown in a non-human species, the pigeon, during a delayed same/different task with color pictures (e.g. Wright et al., in press). Given that the six item colored square set is smaller, there is more repetition across trials resulting in a greater build-up of proactive interference, thus making the task more difficult.

In order to equate the difficulty, and make parameters as similar as possible the monkeys were switched to a task using the same set of six colored squares used to test humans (see Appendix). They were tested with display sizes of two, three, four, five, and six.

This test provided a five point function to which to fit the power law functions, as was done with humans. In addition, three display sizes (two, four, and six) were also tested in humans, thereby permitting a direct comparison between species for these values. These tests provide more direct comparisons with humans given the use of the same viewing times and delays (1000-ms for both viewing time and delay), and testing the monkeys across five display sizes, as was done with humans.

Methods

The subjects were the same two rhesus monkeys used in Chapters 3 and 4. The stimuli and apparatus were also the same as in the previous chapters. All animal procedures were in compliance with National Institute of Health's guidelines and were approved by the University of Texas Health Science Center at Houston Institutional Animal Care and Use Committee.

Testing

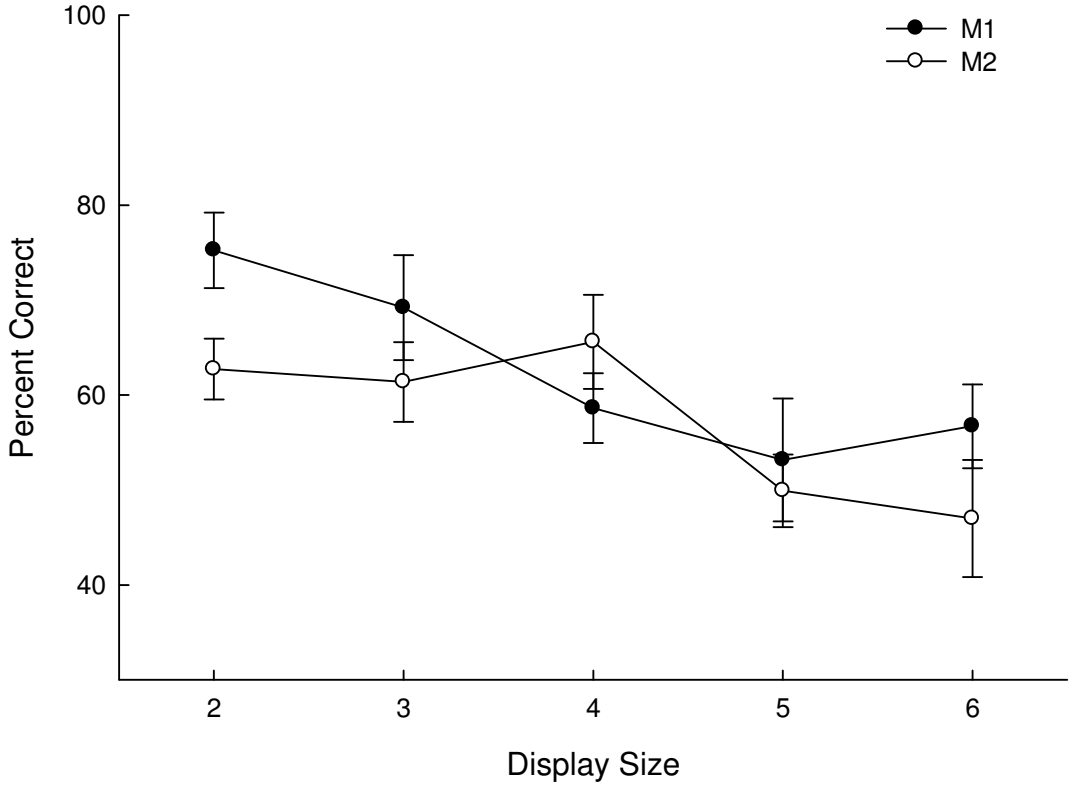
The monkeys were tested with ten alternating 96-trial sessions of both colored squares and clip art. Within each session display sizes of two, three, four, five, and six were intermixed. The delay was always 1000-ms and the viewing time ranged from 1000-5000ms at 500 ms increments in order to provide variability to the monkeys. The variability served to encourage them to maintain vigilance throughout the session. Trials with a viewing time of 1000-ms constituted the majority (approximately 56%) of trials tested in each session and were the only trials included in the analysis presented here. As a result a total of 108 trials per display size were tested across the ten sessions per stimulus type.

Results & Discussion

As shown in Figure 5.1, performance declined as a function of display size for both monkeys. Separate repeated measures ANOVAs of display size × stimulus type showed a significant effect of display size for both monkeys [M1: $F(4, 36) = 11.65$, $p < 0.001$; M2: $F(4,36)$

= 4.33, $p = 0.006$] as well as a significant effect of stimulus type [M1: $F(1,9) = 67.87$, $p < 0.001$; $F(1,9) = 6.80$, $p = 0.03$]. The monkeys performed better with clip art (10.57% difference) than with colored squares.

A



B

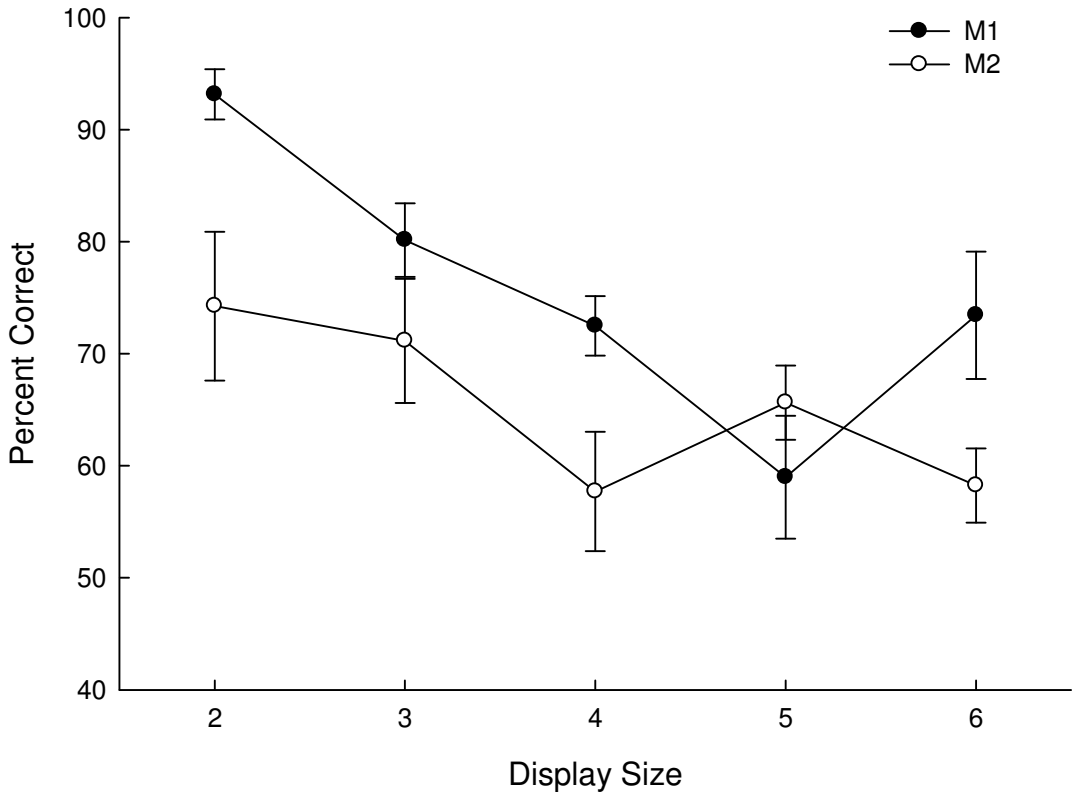


Figure 5.1: Percent Correct in Change Detection Task for A) Colored Squares and B) Clip Art. Error bars represent standard error of the mean.

Estimating capacity. Capacity estimates were calculated using Equation 2.1. Mean capacity estimates for each stimulus type and display size are shown in Figure 5.2. Mean capacity for colored squares was 0.33 ± 0.10 and mean capacity for clip art was 0.84 ± 0.08. Thus, based on a fixed-capacity model of VSTM, monkeys were accurately maintaining less than one stimulus in memory during the delay interval.

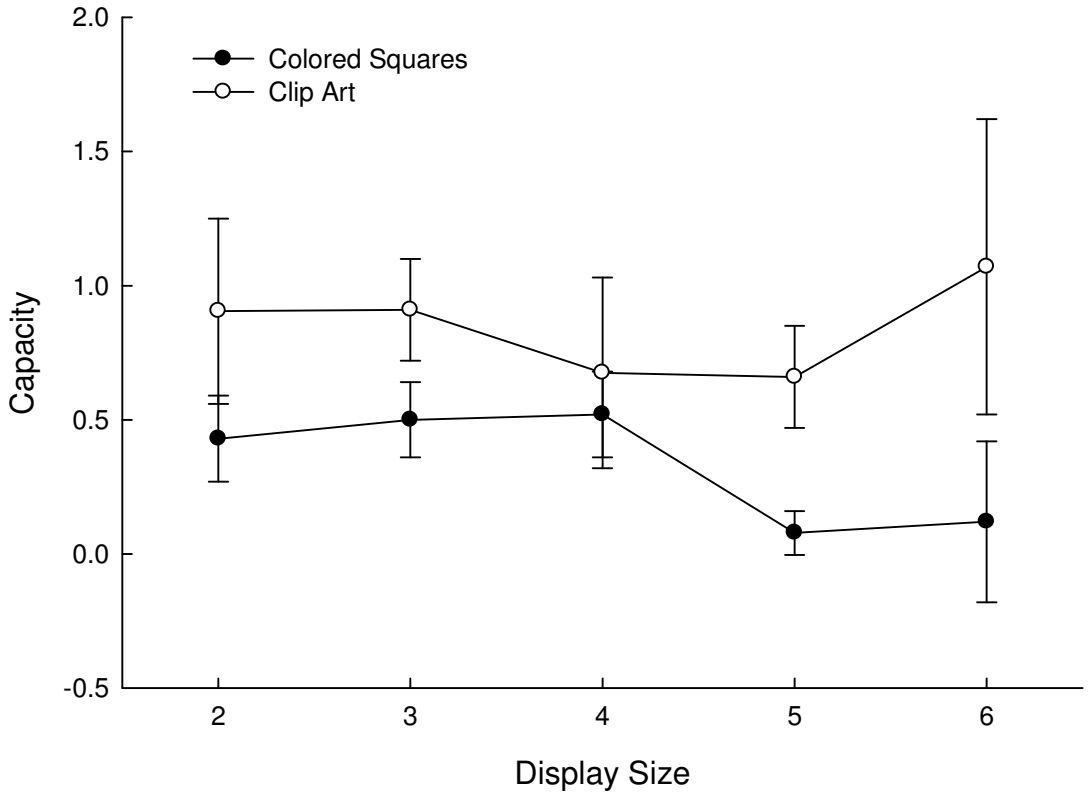
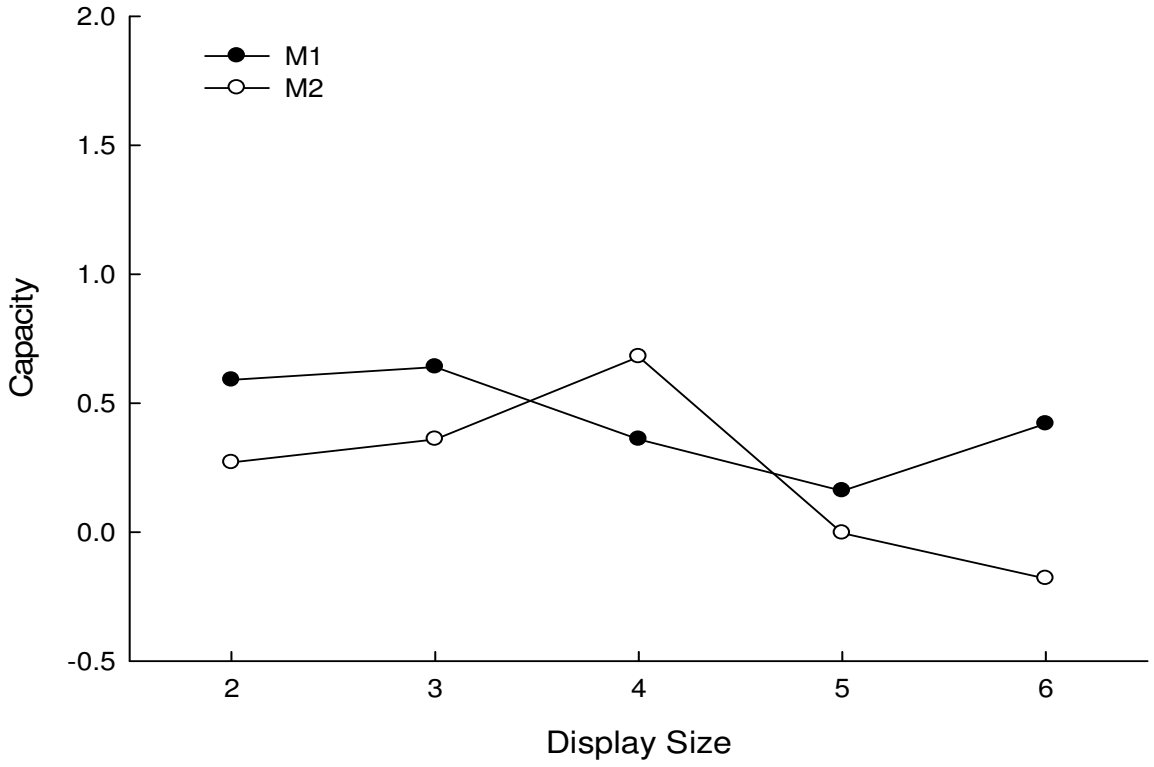


Figure 5.2: VSTM Capacity Estimates for Rhesus Monkeys with Colored Squares & Clip Art. Error bars represent standard error of the mean.

Much like the human subjects in Chapter 2, both monkeys showed more variability in their capacity estimates than would be predicted by a fixed-capacity model of VSTM (Figure 5.3). Following the logic introduced in Chapter 2, that an individual subject's capacity estimates should fall within one standard deviation of their mean, the variability is too great for both subjects with both stimulus types. For instance, with clip art, M1's capacity estimate from the five item display size (0.47) falls below the mean (1.09) by greater than one standard deviation(0.42) whereas his estimate from the six item display (1.62) exceeds the mean by greater than one standard deviation, and the same is true for M2, as his capacity estimate from the five item display size (0.85) exceeds the mean (0.59) of his capacity measures by greater than one standard deviation (0.20). For colored squares, the individual monkeys'

capacity estimates are also shockingly low (M1: 0.43, M2: 0.23), and again show more variability than one would expect. It is unreasonable to believe that they are not adequately storing a single stimulus in memory, particularly given that performance is significantly greater than chance for display sizes two through four for both monkeys with colored squares [single-mean *t*-tests against chance (50%), all $p_s \leq 0.02$]. Put otherwise, how could M1 perform with almost 70% accuracy with the three item display size if he was only accurately maintaining 0.43 stimuli?

A



B

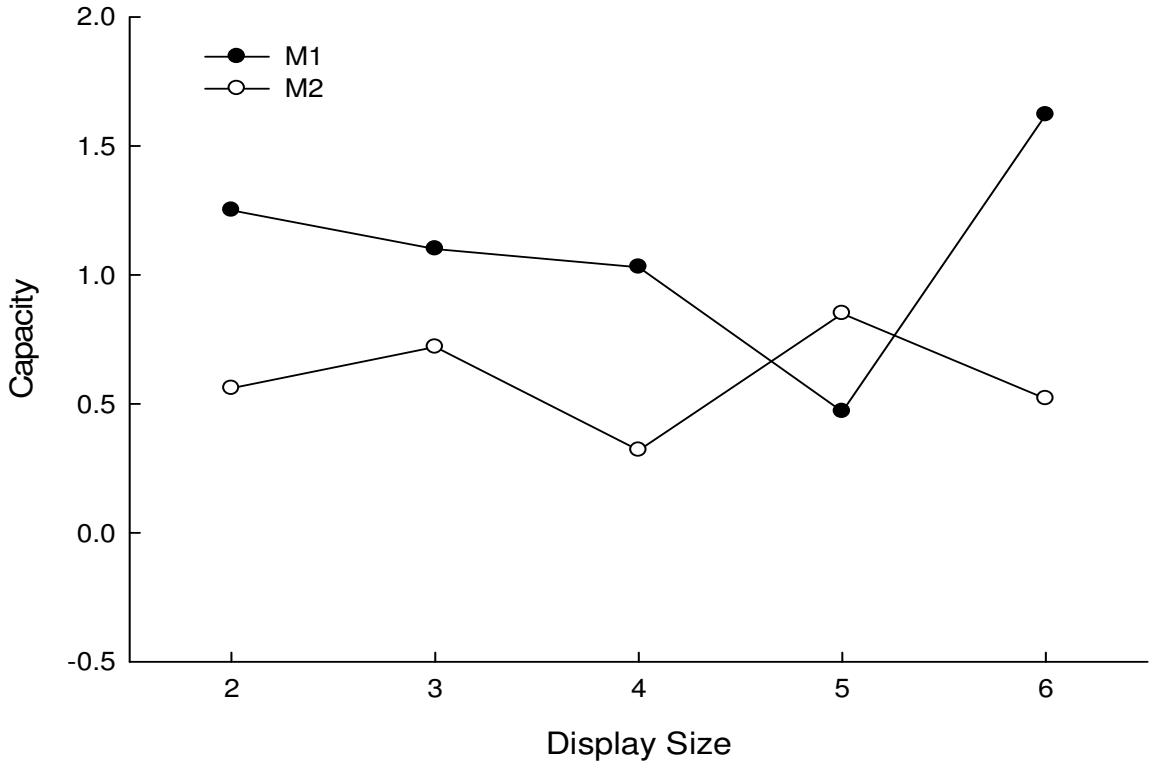


Figure 5.3: Individual Monkey Capacity Estimates for A) Colored Squares and B) Clip Art. Error bars represent standard error of the mean.

Continuous-Resource Model. *d'* values were calculated using Equation 2.2 and are displayed in Figure 5.4. Power law functions were generated as described in Chapter 2. Individual power law fits for clip art had r^2 values of 0.74 and 0.66 for M1 and M2, and for colored squares the r^2 values were 0.84 and 0.72 for M1 and M2 respectively. Furthermore, the mean power law functions for the monkeys provided good fits to the group data with r^2 values of 0.86 and 0.94 for clip art and colored squares respectively (equations provided in legend of Figure 5.4). These r^2 values were statistically significant [Colored Squares: $F(1,8) = 81.80$, $p < 0.0001$; Clip Art: $F(1,8) = 43.57$, $p = 0.0002$].

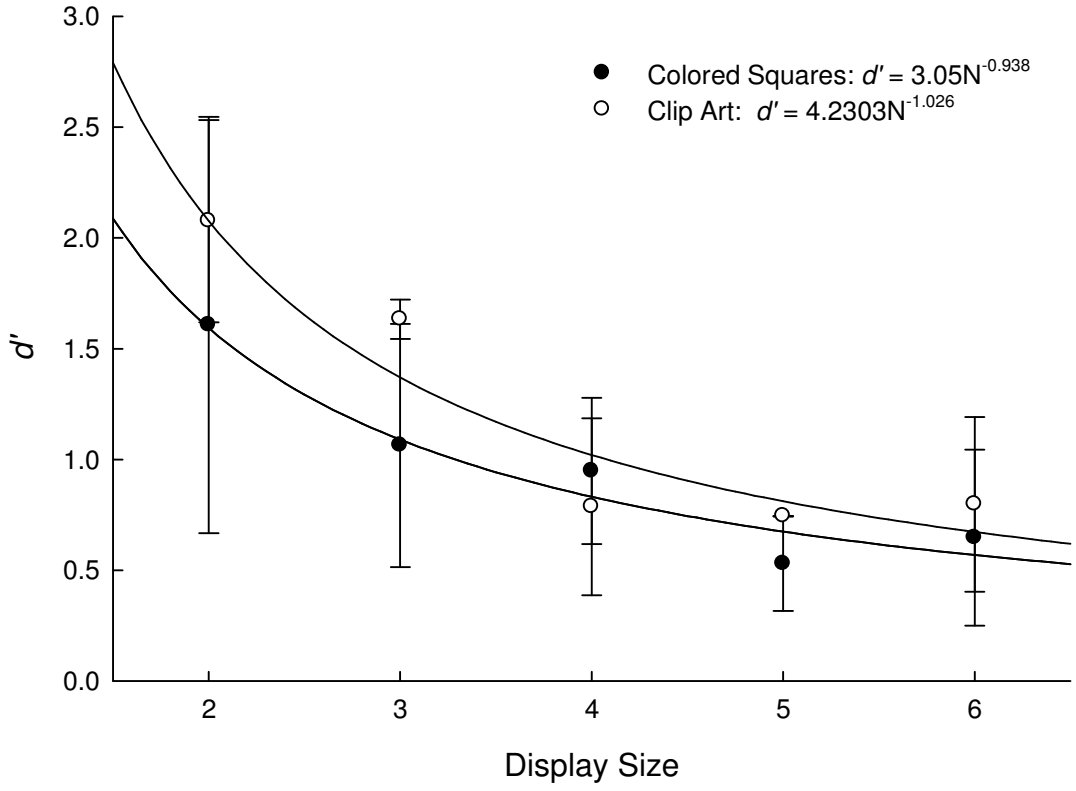


Figure 5.4: Power Law Fits for *d'* for Colored Squares and Clip Art. Error bars represent standard error of the mean.

Comparison to Human Subjects. The data from rhesus monkeys are compared in Figure 5.5 to the data from human subjects in Chapter 2, Experiment 1. Overall performance is compared in Figure 5.5. By comparing the results from the three shared display sizes (two, four and six items), humans outperformed monkeys by 19.24% for clip art trials and by 27.27% for colored square trials. A repeated-measures ANOVA of display size (2, 4, and 6 only) × stimulus type × species revealed a main effect of display size [$F(2,12) = 12.85$, $p = 0.001$] and species [$F(1,6) = 12.13$, $p = 0.01$]. There was also a significant interaction of stimulus type × species [$F(1,6) = 12.18$, $p = 0.01$]. The interaction resulted from the fact that monkeys performed better with clip art than they did with colored squares (10.57% difference), whereas the human subjects did not show a significant difference between these two stimulus types.

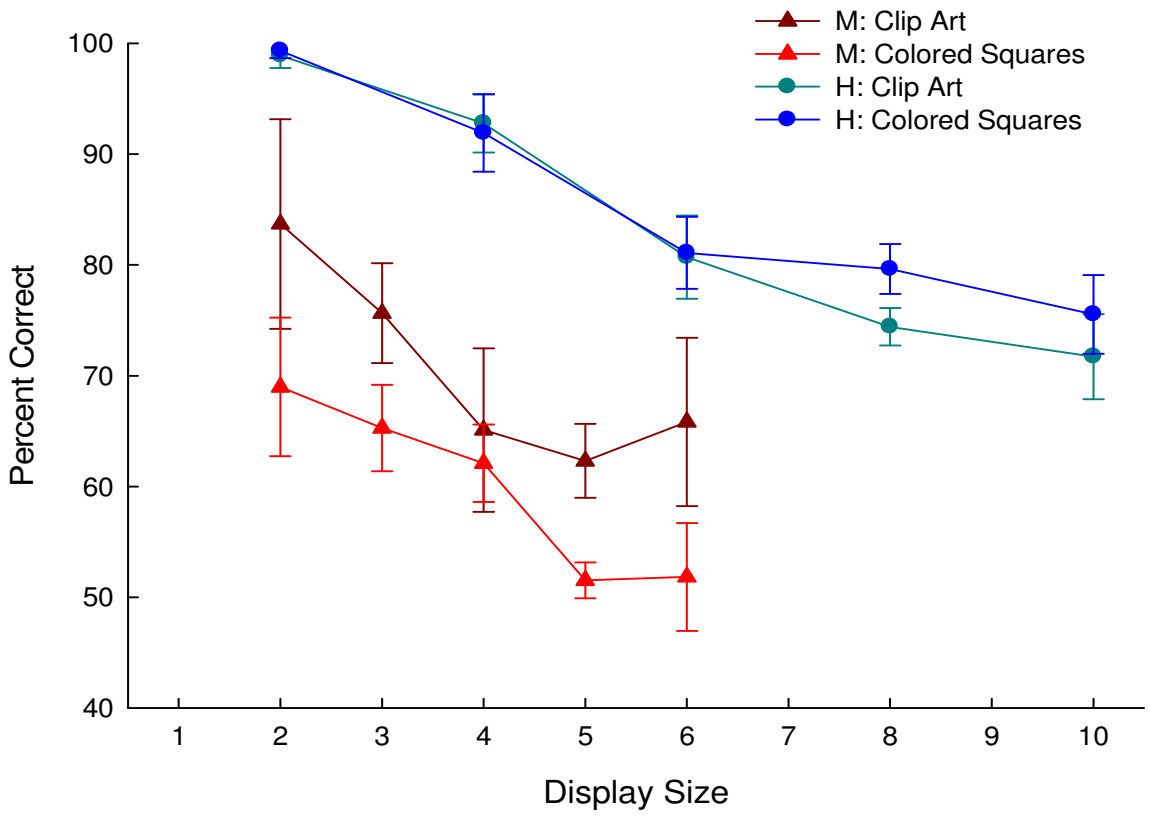


Figure 5.5: Change Detection Performance by Humans & Monkeys. Viewing Times and Delay Intervals are 1-second for both species. Error bars represent standard error of the mean.

Capacity Estimates. Capacity estimates between the two species were also compared (Figure 5.6). Mean capacity estimates for humans were 2.78 ± 0.39 for clip art and 2.46 ± 0.35 for colored squares, whereas mean capacity estimates for monkeys were 0.84 ± 0.08. for clip art and 0.33 ± 0.10 for colors. It should be noted that this is an indirect comparison because humans were tested at display sizes of 2, 4, 6, 8, and 10 whereas monkeys were tested at display sizes of 2, 3, 4, 5, and 6.

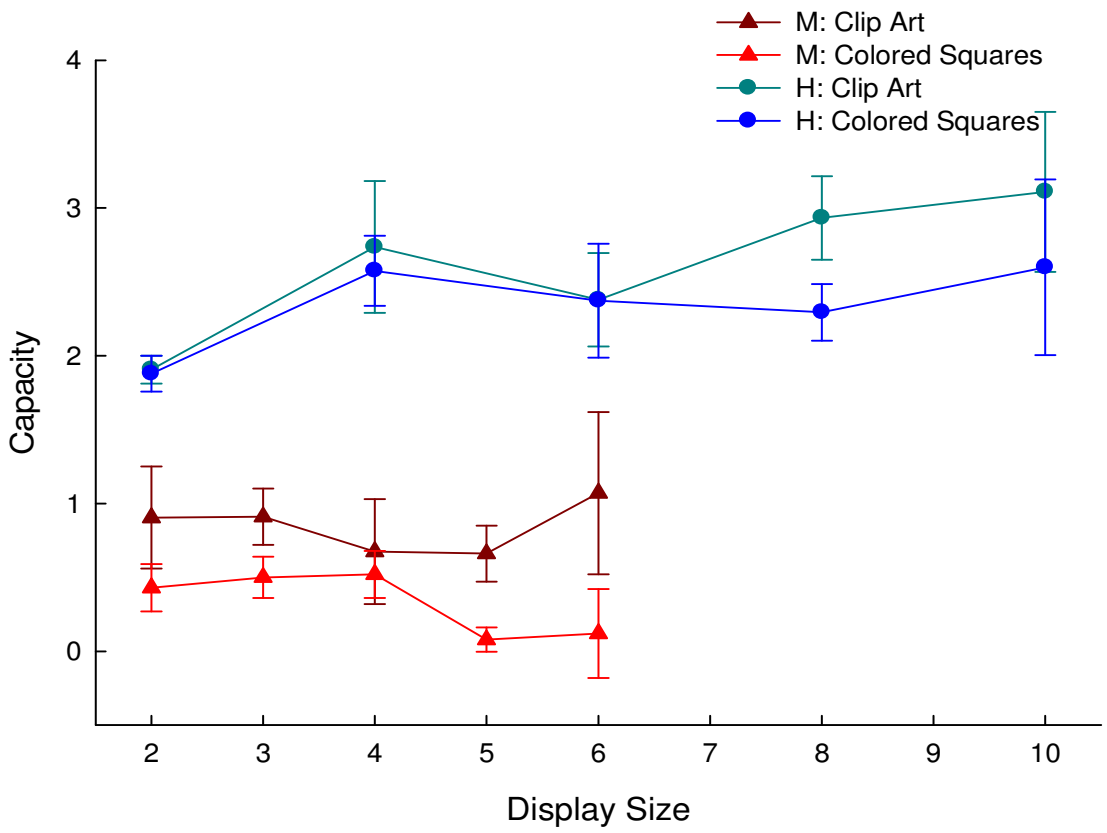


Figure 5.6: Capacity Estimates for Rhesus Monkeys & Humans with Colored Squares & Clip Art. Error bars represent standard error of the mean.

Continuous-Resource Model. Monkey and human performance was also compared using *d'* measures as per the continuous-resource model. *d'* values and power law fits for both species are compared in Figure 5.7. Both species' *d'* values are well characterized by power law functions, and the exponents of these functions fall within a close range and are similar across stimulus types. For colored squares, the exponents were -0.94 and -0.86 for monkeys and humans respectively, and for clip art the exponents were -.70 and -1.026 for monkeys and humans, indicating that the shape of the curves are similar across species. An unpaired *t*-test demonstrated that there were no significant differences in exponent value across species [$t(14) = 1.54$, $p = 0.15$]. Not surprisingly, however, the coefficients of the power

law functions were significantly greater for humans [unpaired *t*-test, $t(14) = 2.26$, $p = 0.04$], indicating that overall memory sensitivity (*d'*) is greater in humans.

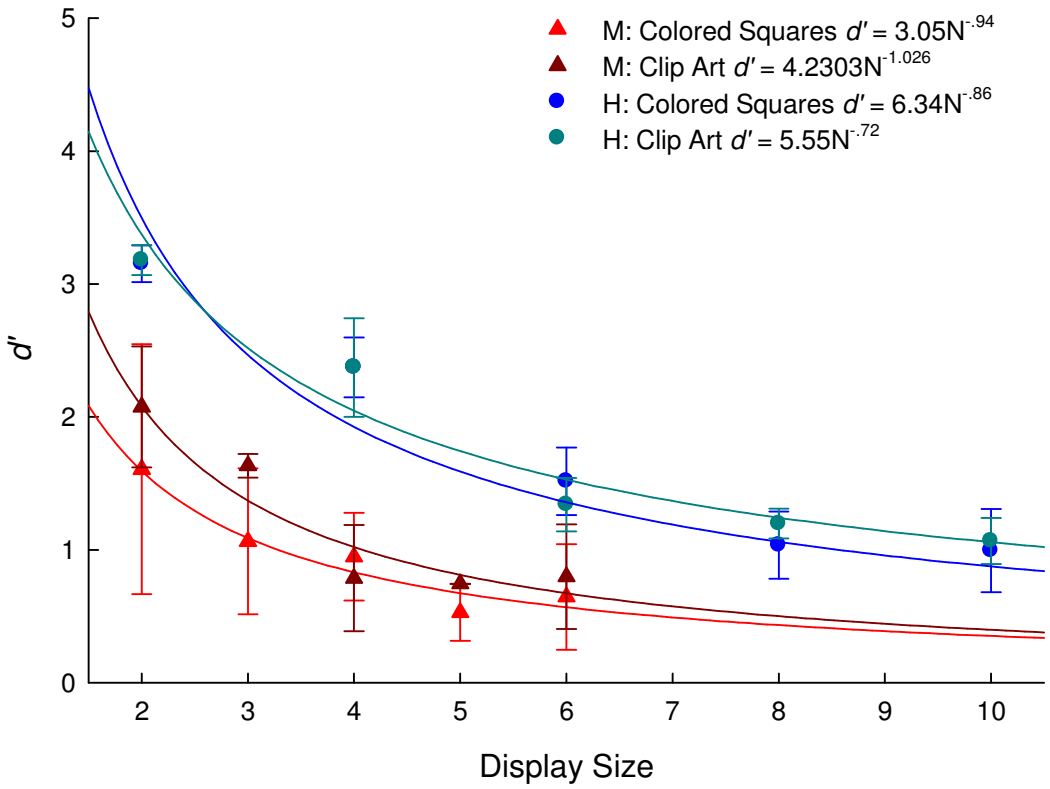


Figure 5.7: Comparison of *d'* Values and Power Law Fits for Humans and Rhesus Monkeys.

Thus, it seems that the continuous-resource model provided a good fit to data from both humans and monkeys, when both species were tested with the same parameters in terms of viewing time and delay.

Meanwhile, the results from the fixed-capacity model are increasingly troubling. The low capacity values obtained for monkeys are difficult to reconcile with prior work with monkeys in list memory (e.g. Wright et al., 2007). But even more troubling are the capacity values in relation to the results obtained here in Chapter 5. How can monkeys perform at above chance levels in this task if they cannot reliably store a single item in VSTM? One might argue that the fixed-capacity model is designed for humans, and thus is not applicable to

rhesus monkeys. However, even within the human literature, some authors report a wide range of capacity estimates from low and high capacity individuals (e.g., Vogel & Machizawa, 2004). They report a low end of the range around 1.5 stimuli which is comparable to the capacity estimates obtained from M1 in the clip art condition. Thus, if a monkey can perform as well as some "low capacity" individuals, then the model should be applicable. Furthermore, given the behavioral similarities in how the monkeys and humans perform the task (looking for change in a general sense – Chapter 3), comparing monkeys and humans using the same model framework is warranted. The fact that the framework of the continuous-resource model accurately predicts a pattern of performance (*d'* values that fall in a power law fashion as a function of display size) both supports the notion that the same model can be used for two species, and that the continuous-resource model is more likely to represent what actually occurs in VSTM.

CHAPTER 6: CONCLUDING REMARKS

In this series of experiments, it has been shown for the first time that rhesus monkeys can be trained and tested in the change detection task such that their VSTM can be directly compared to humans. The monkeys learned the task quickly, performed it using short-term memory (rather than attentional capture), and readily transferred to several novel types of change and thereby demonstrated that they performed the task in a manner similar to humans. They were shown to be looking for change in a general sense, which is similar if not exactly what human subjects are instructed to do.

By comparing monkeys to humans, predictions of two predominant models of VSTM were tested, the fixed-capacity model and the continuous-resource model. In our studies with both humans and rhesus monkeys, we identified several puzzling results that are difficult to justify in the context of the fixed-capacity model. Interestingly, these problematic results (which will be outlined below) were resolved by taking the perspective of the continuous-resource model.

Inconsistent capacity measures across display sizes. In Chapter 2, Experiment 1, subjects were tested with five display sizes. The formula used to estimate capacity (Equation 2.1) takes display size into account in order to compute capacity based on the empirical accuracy at a given display size. Thus, capacity estimates from the various display sizes should roughly agree. However, in Experiment 1, we found that some subjects had highly variable capacity estimates across display sizes. For instance, S2's capacity with colored squares ranged from 2.54 (four item display) to 5.45 (ten item display). Likewise, S6's capacity with clip art ranged from 1.65 (four item display) to 5.29 (ten item display). While those are the two most extreme examples from Experiment 1, other subjects had capacity changes of approximately 1.5 to 2 slots across display sizes which is inconsistent with the notion of a capacity limited slot like storage system. The system should reliably be storing the same amount of information. The same large amount of variability across display sizes was

also seen in Experiment 2 of Chapter 2. Six of seven human subjects showed capacity measures that differed by 1.5 slots or more across display sizes. There were also more extreme individual examples, for instance with Kanji characters, S2's results demonstrated that the subject had slots for approximately 1.62 characters in the six item condition, but had slots for 4.83 stimuli in the ten item condition. The same sort of inconsistency in capacity estimates was identified in Chapter 5 for the monkeys. As such it seems that both species show wildly variable capacity estimates when display size is varied. This conclusion is not compatible with a capacity limited slot-like storage system, because such a system should by definition be consistent across display sizes, or at least vary in a small range consistent with the population variability suggested by the "magic number 4 ± 1". After all, testing subjects at varying display sizes is how capacity itself is estimated. Because the continuous-resource model does not predict a fixed and completely filled slot-like storage system, the finding of capacity variability is not problematic for the continuous-resource model. Also, the continuous-resource model predicts variability in memory resource across trials (within an individual subject) which can help explain the variability found for both human and monkey subjects (Ma et al., under review).

Imperfect performance at display sizes less than capacity. Another prediction of the fixed-capacity model is that performance should be perfect when the display size tested is less than capacity. Because all subjects had capacity estimates greater than two in both the colored square and the clip art conditions, performance should have been perfect by all subjects with two item display sizes. However, S3 was 93.3% accurate with colored squares (indicating a capacity of 1.27) and S4 was 96% accurate with clip art (indicating a capacity of 1.43). Imperfect performance with small display sizes is not a problem for the continuous-resource model. The model postulates that VSTM should be flexibly allocated among stimuli. Flexible allocation is not necessarily optimal, such that on occasion, even with small display

sizes such as two, both stimuli will not be perfectly stored and may be subsequently forgotten or confused.

Performance differences based on stimulus type. While Experiment 1 of Chapter 2 showed that humans performed comparably with both colored squares and clip art (no significant differences in performance), Experiment 2 showed that performance depended on the type of stimulus tested (Snodgrass and clip art were better than Kanji and kaleidoscopes). This result rejects the more rigid interpretation of the fixed-capacity model (e.g. Luck & Vogel, 1997; Cowan, 2001; 2005) which states that capacity should not depend on the type of stimulus. Memory slots can be filled by various types of stimuli, the only limitation is the number of slots available. This model is incompatible with our results, and the flexible approaches to the fixed-capacity model do not provide very satisfactory solutions. One flexible approach is that the fixed-capacity model should be modified by a two component limitation on VSTM (Alvarez & Cavanagh, 2004). One proposed limitation is that the maximum number of stimuli stored had to be four or five. The other proposed limitation was that visual information should vary jointly as a function of both the number of stimuli and the amount of visual information per stimulus. Another proposal was that the variance in performance across stimulus types could be offset by increasing the viewing time for difficult stimuli (Eng et al., 2005). This proposal is unlikely to be a viable solution because naturally increasing the viewing time would improve performance, which in turn should be true for all types of stimuli.

It should be noted that the monkeys also showed a significant effect of stimulus type in Chapter 5. However, the situation is a bit more complex, as they performed worse with colored squares than they did with clip art. The lower performance with colored squares could have been the result of the repetition of colors within and across trials rather than the actual differences in the types of stimuli (colored squares vs. clip art). Repetition across trials leads to increases in proactive interference, and repetition within trials require the subjects to remember both what colors were present as well as where they were located in order to

perform accurately. It is likely that the increases in difficulty associated with the six item color set can explain the differences in performance found in Chapter 5 because the monkeys showed little difference in performance between colors and clip art in Chapter 4 (much like humans).

Monkey capacity of one. The finding that VSTM capacity in monkeys is approximately one item or less is very difficult to reconcile with earlier findings from rhesus monkeys in list-memory tasks. In visual list-memory tasks, rhesus monkeys can perform the task with lists of four or more stimuli with performance levels of 90% correct or better (Wright, 2007). Although performance at serial positions varies based on the delay, the shifts from recency to primacy effects as the delay lengthens would not develop if VSTM capacity were only one. The comparison between change detection and list-memory tasks is somewhat indirect because in change detection stimuli are presented simultaneously, whereas in list memory stimuli are presented sequentially. This can perhaps explain why overall performance levels differ between list memory and change detection. While performance with four-item lists has been shown to be 90% or greater, in change detection the monkeys' performance was about 70% correct with four-item displays. This overall difference in performance can be explained using the continuous-resource model. The simultaneous presentations used in change detection require the monkey to optimally divide his attention across space among all the stimuli in the sample. In list memory, the stimuli are presented one at a time in a fixed location such that the monkey only needs to allocate his attention to one area of space. Suboptimal allocation of attention in change detection could result in noisier representations in VSTM. Increased noise in the memory representation would result in lower performance, as predicted by the continuous-resource model.

Color confusion. Due to the large amount of trials collected from the rhesus monkeys, we were able to perform a multi-dimensional scaling analysis to assess the degree to which the monkeys made mistakes when one color changed to another. From this analysis we

discovered that monkeys routinely made mistakes when one color changed to a similar color. For instance, M1 frequently confused purple and magenta (his performance was 52% correct when one changed to the other). Likewise, M2 frequently confused purple and blue (his performance was 57% correct when one changed to the other). This result is incompatible with the fixed-capacity model's assertion that VSTM is a high resolution, noise free storage system. The fixed-capacity model states that an item should be stored perfectly (within the capacity limits) or not stored at all. Thus, it is difficult to reconcile the finding of color confusion within the context of the fixed-capacity model. However, the continuous-resource model can easily account for the finding of color confusion. The model predicts that stimuli are represented in memory with noise. Noisy representations of similar colors should be easily confused, as the monkeys were found to do.

The continuous-resource model. Aside from its ability to make sense of the problematic findings described above, the continuous-resource model also provides a good fit to the data, generally. In all cases, with all types of stimuli, for both humans and monkeys, *d'* values were extremely well fit by power law functions. The main predictions of the continuous-resource model are that memory sensitivity (*d'*) should decline with display size and that the decline should be well fit by a power law function. These predictions were confirmed in all cases and are thus consistent with the continuous-resource model. Furthermore, the continuous-resource model can reconcile many of the problems identified with a fixed-capacity account of VSTM. Moreover, the continuous-resource model is tied to what is known about computations in the nervous system. The prediction of noisy memory representations is consistent with the physiological properties of the brain (Faisal et al.,2008).

Comparing VSTM between monkeys and humans. Qualitative similarities between humans and monkeys are apparent in the performance comparisons in Chapters 4 and 5, as they both show a decline in performance as display size increases. There was a quantitative difference between species as humans outperformed monkeys by an average of 19.25% in

Chapter 4 and 23.25% in Chapter 5. However, once viewed in the context of the continuous-resource model, it is striking how qualitatively similar VSTM is between the two species. Both species' memory sensitivity (*d'*) values were well fit by power law functions and the power law functions even have similar exponents. The qualitative similarity between species was expected based on what is known about brain regions involved in VSTM in both species, including visual cortex and the prefrontal cortex. However, the great advantage of these between-species comparisons is that they provide converging evidence in favor of the continuous-resource model. Thus, by testing rhesus monkeys with the same procedures and stimuli used with humans, we were able to gain a better understanding of human cognition by identifying the continuous-resource model as the more plausible account of VSTM function.

Future directions. Establishing the rhesus monkey as an animal model for VSTM that can be tested with the same procedures as humans lays the foundation for future work investigating the neurobiological basis of VSTM. Future studies using invasive techniques can be conducted to further investigate the brain areas that subserve VSTM, and how the regions work together in a unified network. All neurobiological investigations of VSTM to date using change detection have been guided by the notion of a fixed capacity (e.g., Todd & Marois, 2004; Vogel & Machizawa, 2004; Buschman et al., 2011). All of these studies interpreted their results as proof for a capacity-limited storage system. However, given Wilken & Ma's contention that the "magic number 4±1" is an artifact of the noise which increases as display size increases, the neurobiological evidence identified in these studies may be biased by this artifact, and therefore misinterpreted.

The continuous-resource model most likely provides a more plausible model framework and future studies should consider this framework as a guide to neurobiological investigations. In fact, neural investigations of VSTM may be useful in providing further tests of the continuous-resource model. Ma et al. (under review) have suggested that VSTM is roughly equivalent to attentional gain. Specifically, they predict that neural gain is associated

with the encoding process, and that this gain varies across trials, but generally decreases as display size increases. A neurophysiological investigation of neural gain in the context of a VSTM experiment would thus provide further support for, or potentially refute the predictions of the continuous-resource model.

To conclude, this dissertation demonstrates that combining tools from comparative psychology, cognitive neuroscience and computational neuroscience can provide a more complete understanding of one functions of the brain, VSTM. Combining the power of these three fields (and others) in the future will undoubtedly provide great insight into the mechanisms of cognition and behavior.

APPENDIX

Procedural Note Regarding Chapter 5:

The monkeys were first tested with display sizes of two, four, six, eight, and ten with both colored squares and clip art (just like humans). This task proved to be difficult for the monkeys, as performance at six, eight, and ten was at or near chance. This made the fits of the continuous-resource model's power law functions rather poor, since *d'* values were at or very near zero for the three largest display sizes, resulting in a non-curvilinear function. In addition, the capacity estimates from these three display sizes were at or near zero, which is difficult to interpret.

Although it is important to test VSTM at the limits of the individual's ability, given that performance was at or near chance at three of the five display sizes tested and hence three-fifths of trials, it is likely that this test was too difficult for the monkeys and does not provide an ideal assessment of their VSTM abilities. The difficulty of the test may have also hurt the monkeys' motivation to perform the task since they were only receiving reinforcement on approximately 60% of trials at best. Thus, the monkeys were retested with display sizes of two, three, four, five and six to improve performance while still obtaining a five point function and enabling a direct comparison with humans at the display sizes of two, four, and six.

BIBLIOGRAPHY

Alvarez, G.A., & Cavanagh, P. (2004). The Capacity of Visual Short-Term Memory is Set Both by Visual Information Load and by Number of Objects. *Psychological Science, 15,* 106-111.

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.) *The psychology of learning and motivation.* Vol. 2, (pp.89-105). New York: Academic Press.

Baddeley, A. D. (1986). *Working memory.* Clarendon Press.

Baddeley, A., Bressi, S., Della Sala, S., Logie, R., & Spinnler, H. (1991). The Decline of Working Memory in Alzheimer's Disease: A Longitudinal Study. *Brain, 114,* 2521-2542.

Baddeley, A., Logie, R., Bressi, S., Della Sala, S., & Spinnler, H. (1986). Dementia and Working Memory. *The Quarterly Journal of Experimental Psychology, 38A,* 603-618.

Bays, P.M., Catalao, R.F.G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9,* 1-11.

Bays, P.M. & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321,* 851.

Beck, J.M., Ma, W.J., Kiani, R., Hanks, T.D., Churchland, A.K., Roitman, J.D., Shadlen, M.N., Latham, P.E., and Pouget, A. (2008). Bayesian decision-making with probabilistic population codes. *Neuron, 60,* 1142-1145.

Bolhuis, J.J., & van Kampen, H.S. (1988). Serial position curves in spatial memory of rats: primacy and recency effects. *The Quarterly Journal of Experimental Psychology, 40,* 135-149.

Brandes, D., Ben-Schachar, G., Gilboa, A., Bonne, O., Freedman, S., & Shalev, A.Y. (2002). PTSD symptoms and cognitive performance in recent trauma survivors. *Psychiatry Research, 110,* 231-238.

Buchanan, J.P., Gill, T.V., & Braggio, J.T. (1981). Serial position and clustering effects in chimpanzee's "free recall." *Memory & Cognition, 9,* 651-660.

Budson, A.E., & Price, B.H. (2005). Memory Dysfunction. *The New England Journal of Medicine, 352,* 692-699.

Buschman, T.J., Siegel, M., Roy, J.E., & Miller, E.K. (2011). Neural substrates of cognitive capacity limitations. *Proceedings of the National Academy of Sciences, 108,* 11252-5.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104,* 163-191.

Cowan, N. (1995). *Attention and memory: An integrated framework.* Oxford Psychology Series, No. 26. Oxford University Press.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87-185.

Cowan, N. (2005). *Working Memory Capacity* (New York: Psychology Press).

Darwin, C. (1872). *The Expression of Emotions in Man and Animals*. University of Chicago Press., Chicago, IL.

Darwin, C.J., Turvey, M.T., & Crowder, R.G. (1972). An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology, 3,* 255-267.

Deaner, R.O., Isler, K., Burkart, J., & van Schaik, C. (2007). Overall Brain Size, and Not Encephalization Quotient, Best Predicts Cognitive Ability across Non-Human Primates. *Brain, Behavior, & Evolution, 70,* 115-124.

Dubois, B. & Pillon, B. (1997). Cognitive deficits in Parkinson's disease. *The Journal of Neurology, 244,* 2-8.

Elmore, L.C., Wright, A.A., Rivera, J.J., & Katz, J.S. (2009). Individual differences: either relational or item-specific learning in a *same/different* task. *Learning & Behavior, 37,* 204-213.

Elston, G.N., Benavides-Piccione, R., and DeFelipe, J. (2001). The Pyramidal Cell in Cognition: A Comparative Study in Human and Monkey. *The Journal of Neuroscience, 21,* 1-5.

Eng, H.Y., Chen, D. & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review, 12,* 1127-1133.

Faisal, A.A., Selen, L.P.J., & Wolpert, D.M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience, 9,* 292-303.

Fukuda, K., & Vogel, E.K. (2010). Visual short term memory serves as a gateway to long term memory. *Journal of Vision,10.*

Fukuda, K., & Vogel, E.K. (2011). Visual short term memory also gates long term memory without explicit retrieval. *Journal of Vision, 11.*

Funahashi, S., Bruce, C.J., & Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology, 61,* 331-349.

Fuster, J.M., & Bauer, R.H. (1974). Visual short-term memory deficit from hypothermia of frontal cortex. *Brain Research, 81,* 393-400.

Gabrieli, J.D.E. (1998). Cognitive Neuroscience of Human Memory. *Annual Review of Psychology, 49,* 87-115.

Green, D.M., & Swets, J.A. (1966). Signal Detection THoery and Psychophysics. (New York: Wiley).

Harper, D.N., McLean, A.P., & Dalyrymple-Alford, J.C. (1993). List item memory in rats: Effects of delay and delay task. *Journal of Experimental Psychology: Animal Behavior Processes, 19,* 307-316.

Heyselaar, E., Johnston, K., & Paré, M. (2011). A change detection approach to study visual working memory of the macaque monkey. *Journal of Vision, 11,* 1-10.

Katz, J.S., Wright, A.A., & Bachevalier, J. (2002). Mechanisms of *same/different* abstract-concept learning by rhesus monkeys *(Macaca mulatta). Journal of Experimental Psychology: Animal Behavior Processes, 28,* 358-368.

Kesner, R.P., & Novak, J.M. (1983). Serial position curve in rats: role of the dorsal hippocampus. *Science, 218,* 173-175.

Koenen, K.C., Driver, K.L., Oscar-Berman, M., Wolfe, J., Folsom, S., Huang, M.T., & Schlesinger, L. (2001). Measures of Prefrontal System Dysfunction in Posttraumatic Stress Disorder. *Brain and Cognition, 45,* 64-78.

Kumar, S., & Hedges, S.B. (1998). A molecular timescale for vertebrate evolution. *Nature, 392,* 917-920.

Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279-281.

Ma. W.J., Beck, J.M., Latham, P.E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience, 9,* 1432-1438.

Ma, W.J., Beck, J.M., & Pouget, A. (2008). Spiking networks for Bayesian inference and choice. *Current Opinion in Neurobiology, 18,* 217-222.

Ma, W.J., Keshvar, S., & van den Berg, R. (under review). Change detection as probabilistic inference under variable resources.

Macmillan, N.A., and Creelman, C.D. (2005). Detection Theory: A User's Guide, Second Edition (Mahwah, NJ: Lawrence Erlbaum Associates).

Mishkin, M. & Delacour, J. (1975). An analysis of short-term memory in the monkey. *Journal of Experimental Psychology: Animal Beahvior Processes, 1,* 326-334.

Pashler, H., (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44,* 369-378.

Petrides, M. (1994). Frontal lobes and working memory: evidence from investigations of the effects of cortical excisions in nonhuman primates. In *Handbook of neuropsychology* (ed. F. Boller & J. Grafman), vol. 9, pp. 59-82.

Petrides, M. (1996). Specialized systems for the processing of mnemonic information within the primate frontal cortex. *Philosophical Transactions of the Royal Society of London: B Biological Sciences, 351,* 1455-1462.

Petrides, M. (2005). Lateral prefrontal cortex: architectonic and functional organization. *Philosophical Transactions of the Royal Society of London: B Biological Sciences, 360,* 781-795.

Pessoa, L. & Ungerleider, L.G. (2004). Neural correlates of change detection and change blindness in a working memory task. *Cerebral Cortex, 14,* 511-520.

Reed, P. Croft, H., & Yeomans, M. (1996). Rats' memory for serially presented novel flavours: Evidence for non-spatial primacy effects. *The Quarterly Journal of Experimental Psychology, 49B,* 174-187.

Rensink, R.A. (2002). Change Detection. *Annual Review of Psychology, 53,* 245-277.

Rhesus Macaque Genome Sequencing and Analysis Consortium. (2007). Evolutionary and biomedical insights from the rhesus macaque genome. *Science, 316,* 222-234.

Roberts, W.A., & Kraemer, P.J. (1981). Recognition memory for lists of visual stimuli in monkeys and humans. *Animal Learning & Behavior, 9,* 587-594.

Roth, G., & Dicke, U. (2005). Evolution of the brain and intelligence. *TRENDS in Cognitive Sciences, 9,* 250-257.

Rouder, J.N., Morey, R.D., Cowan, N., Zwilling, C.E., Morey, C.C., & Pratte, M.S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences, 105,* 5975-5979.

Sala, J.B., & Courtney, S.M. (2007). Binding of what and where during working memory maintenance. *Cortex, 43,* 5-21.

Sawaguchi, T., & Goldman-Rakic, P.S. (1991). D1 dopamine receptors in prefrontal cortex: involvement in working memory. *Science, 251,* 947-950.

Semendeferi, K., Lu., A., Schenker, N., & Damasio, H. (2002). Humans and great apes share a large frontal cortex. *Nature Neuroscience, 5,* 272-276.

Shepard, R.N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior, 6,* 156-163.

Sperling, G. (1963). A model for visual memory tasks. *Human Factors, 5,* 19-31.

Sperling, G. (1960) The information available in brief visual presentations. *Psychological Monographs, 74,* 498.

Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, 25,* 207-222.

Standing, L., Conezio, J., & Haber, R.N. (1970). Perception and memory for pictures: Single trial learning of 2560 visual stimuli. *Psychonomic Science, 19,* 169-179.

Todd, J.J., & Marios, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428,* 751-754.

Van Essen, D. (2004). Organization of visual areas in macaque and human cerebral cortex. In *The Visual Neurosciences* (Vol. 1), (Chalupa, L.M. and Werner, J.S., eds), pp. 507-521, MIT Press.

Vogel, E.K., & Machizawa, M.G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428,* 748-751.

Vogel, E.K., Woodman, G.F., & Luck, S.J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 92-114.

Voss, J.L. (2009). Long-term associative memory capacity in man. *Psychonomic Bulletin & Review, 16,* 1076-1081.

Wilson, F.A., Scalaidhe, S.P., & Goldman-Rakic, P.S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science, 260,* 1955-1958.

Wright, A.A. (1999). Visual list memory in capuchin monkeys *(Cebus paella*). *Journal of Comparative Psychology, 113,* 74-80.

Wright, A.A. (2007). An experimental analysis of memory processing. *Journal of the Experimental Analysis of Behavior, 88,* 405-433.

Wright, A.A., & Katz, J.S. (2006). Mechanisms of *same/different* concept learning in primates and avians. *Behavioral Processes,* 72, 234-254.

Wright, A.A., Katz, J.S., & Ma, W.J. (in press). How to be proactive about interference: lessons from animal memory. *Psychological Science.*

Wright, A.A., Santiago, H.C., Sands, S.F., Kendrick, D.F., & Cook, R.G. (1985). Memory Processing of serial lists by pigeons, monkeys and people. *Science, 229,* 287-289.

Wright, A.A., Rivera, J.J., Katz, J.S., & Bachevalier, J. (2003). Abstract-Concept Learning and List-Memory Processing by Capuchin and Rhesus Monkeys. *Journal of Experimental Psychology: Animal Behavior Processes, 29,* 184-198.

Zhang, W., & Luck, S.J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453,* 233-235.

VITA

Lauren Caitlin Elmore was born in Round Rock, Texas on June 23, 1984, the daughter of Carol Jean Elmore and Robert James Elmore. She has one older sister, Christa Nicole Parker, who lives in Harlingen, TX, and an older brother, Christopher Marcus Elmore who resides in Katy, TX. She attended high school at the Madeira School in McLean, Virginia and graduated in May 2002. She then entered the University of Texas at Austin, where she earned a Bachelor of Science Degree in Psychology with Special Honors in May of 2006. In August 2006, she entered the Ph.D. program in Neuroscience at the University of Texas at Houston Graduate School of Biomedical Sciences.

Permanent address:

1711 Old Spanish Trail #248

Houston, TX 77054