The Influence of Gestalt Principles of Grouping on Visual Working Memory Representations

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Psychology

by

Dwight James Peterson

Dr. Marian E. Berryhill/Dissertation Advisor

May 2014
We recommend that the dissertation prepared under our supervision by

DWIGHT J PETERSON

entitled

The Influence Of Gestalt Principles Of Grouping On Visual Working Memory Representations

be accepted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Dr. Marian Berryhill, Advisor

Dr. Gideon Caplovitz, Committee Member

Dr. Michael Crognale, Committee Member

Dr. George Bebis, Committee Member

Dr. Thomas Nickles, Graduate School Representative

Marsha H. Read, Ph.D., Dean, Graduate School

May, 2014
Abstract

Recent studies have demonstrated that bottom-up factors, such as Gestalt grouping cues, can influence the storage of information within visual working memory (VWM). However, the neural correlates underlying grouping-related benefits to VWM performance remain unclear. In the present experiments we introduced and manipulated Gestalt principles (Experiment 1: similarity and proximity) and related grouping cues (Experiments 2 and 3: uniform connectedness) in memory arrays presented during a VWM color change detection task. We monitored the number of representations being stored within VWM during the delay period by measuring the contralateral delay activity (CDA). In the current experiments, we observed both the presence (Experiment 1) and absence (Experiments 2 and 3) of grouped-related benefits to VWM performance. Grouping-related VWM performance benefits were accompanied by a reduction in CDA amplitude. However, when grouping-related benefits were absent, no reduction in CDA amplitude was apparent. The current findings indicate that, when grouping cues are effective, fewer neural resources are required to maintain grouped relative to ungrouped stimuli within VWM.
I dedicate this dissertation to a source of inspiration I may never fully comprehend.

Inspiration, move me brightly.
Acknowledgements

The number of people and sources of inspiration requiring acknowledgment pertaining to this accomplishment are vast. First and foremost, I am eternally grateful to my fiancé, Izzy, for her love, support, and patience throughout this process. Additionally, all my love and gratitude are extended to my tribe of friends, family, and especially to my parents, Scott and Sandy. Without their teachings and loving support throughout my life I would have never made it this far.

I am extremely grateful to my mentors at the University of Northern Iowa, Dr. Kim MacLin and Dr. Otto MacLin. They saw the sparkle in my eye while teaching me about human cognition and watched my initial curiosities evolve into a transitive nightfall of diamonds. Perhaps most importantly, they helped me to free and open my mind to a world of possibility.

My academic life here at UNR started with a great deal of uncertainty. Without a net during my first year, I wondered where I would land. In my initial conversations with Dr. Marian Berryhill during my second year, it became very apparent that the opportunity to take my scientific training to the next level was there for the taking. From her mentoring, guidance, and ability to provide ample funding (e.g., NEI R15EY022775, NIH 1P20GM103650-01, and faculty start-up funding), I have attained an advanced understanding and appreciation of the complexity inherent to the brain. I am grateful to Dr. Gideon Caplovitz, who taught me a variety of advanced technical and problem solving skills that have allowed my abilities as a researcher to evolve at an exponential rate. Finally, I am grateful to my committee members, Dr. Michael Crognale, Dr. Thomas Nickles, and Dr. George Bebis for their time and helpful comments throughout the
dissertation process. My academic family at UNR also deserves acknowledgment, as they have allowed me to keep my sanity by being there with me during both moments of extreme joy and frustration. Thanks especially to my brother in experimentation, Dr. Dan “Trey” McCarthy, who is truly the kindest of the kind and one of the most impressive human beings I have ever had the privilege of knowing. The graduate school experience has been infinitely more enjoyable with you at my side. Thanks to my lab mate, Dr. Kevin Jones, for not taking life too seriously, for making me aware of the superiority of the Pacific Northwest and for still speaking to me after sharing countless conference hotel rooms, cab rides, plane rides, train rides, and a seven by ten foot human hamster cage as an office with me for 4 years. To all of my other brothers and sisters I have met along my journey: Thank you. You know who you are and where you have been.

Finally, I must acknowledge the people responsible for the sources of inspiration that I have acquired throughout my academic life. At the very least, the following people provided me with something amazing to listen to while completing this dissertation. Many thanks to The Grateful Dead, The String Cheese Incident, Phish, Lotus, STS9, Railroad Earth, Widespread Panic and the myriad of other sources of musical ecstasy that have kept a smile on my face during this process. Many thanks to the good folks at the Reno Homebrewer for the supplies, Great Basin, Russian River, Sierra Nevada, Ninkasi, Magnolia, Mad River, Boneyard, Lagunitas, North Coast, and the numerous other sources of liquid libation which I have consumed with great gusto throughout the past five years. Finally, to Dr. Alexander Shulgin and Dr. Albert Hofmann: Thanks for putting in the time and hard work. Your scientific discoveries have had a profound impact on my cognition.
# Table of Contents

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
</tr>
<tr>
<td>Dedication</td>
</tr>
<tr>
<td>Acknowledgements</td>
</tr>
<tr>
<td>Table of Contents</td>
</tr>
<tr>
<td>List of Tables</td>
</tr>
<tr>
<td>List of Figures</td>
</tr>
</tbody>
</table>

## Chapters

### I. General Introduction

- Working Memory 2
- Visual Working Memory Capacity Limitations 4
- The Unit of Storage Debate 7
- The Structure of Visual Working Memory Debate 9
- Neural Correlates of Visual Working Memory Capacity 13
- Electrophysiological Correlates of Capacity 14
- Gestalt Principles of Grouping 21
- Gestalt Principles in Visual Perception Research 23
- Gestalt Principles in Visual Working Memory Research 26
- Outstanding Questions and Overview of Present Experiments 32

### II. Experiment 1

- Introduction 34
- Hypotheses 34
Method 35
Behavioral Analyses 39
Electrophysiological Recording, Processing, and Analyses 39
Behavioral Results 42
Electrophysiological Results 44
Discussion 47

III. Experiment 2

Introduction 51
Hypotheses 52
Method 52
Behavioral Analyses 54
Electrophysiological Recording, Processing, and Analyses 54
Behavioral Results 55
Electrophysiological Results 57
Discussion 60

IV. Experiment 3

Introduction 64
Hypotheses 64
Method 66
Behavioral Analyses 68
Electrophysiological Recording, Processing, and Analyses 68
Behavioral Results 68
Electrophysiological Results 72
V. General Discussion

Overview of the Current Experiments 78
Implications of the Current Findings 78
Limitations of the Current Experiments 90
Conclusions 92

References 94
List of Tables

Table 1. Experiment 1 Means (standard deviations)  47
Table 2. Experiment 2 Means (standard deviations)  60
Table 3. Experiment 2 Means (standard deviations)  74
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.</td>
<td>An Illustration of Gestalt Principles of Grouping</td>
<td>23</td>
</tr>
<tr>
<td>Figure 2.</td>
<td>Experiment 1 Task Paradigm and Stimulus Configurations</td>
<td>36</td>
</tr>
<tr>
<td>Figure 3.</td>
<td>Experiment 1 Behavioral Results</td>
<td>45</td>
</tr>
<tr>
<td>Figure 4.</td>
<td>Experiment 1 Electrophysiological Results</td>
<td>46</td>
</tr>
<tr>
<td>Figure 5.</td>
<td>Experiment 2 Task Paradigm and Stimulus Configurations</td>
<td>54</td>
</tr>
<tr>
<td>Figure 6.</td>
<td>Experiment 2 Behavioral Results</td>
<td>57</td>
</tr>
<tr>
<td>Figure 7.</td>
<td>Experiment 2 Electrophysiological Results</td>
<td>59</td>
</tr>
<tr>
<td>Figure 8.</td>
<td>Experiment 3 Task Paradigm and Stimulus Configurations</td>
<td>67</td>
</tr>
<tr>
<td>Figure 9.</td>
<td>Experiment 2 Behavioral Results</td>
<td>71</td>
</tr>
<tr>
<td>Figure 10.</td>
<td>Experiment 3 Electrophysiological Results</td>
<td>74</td>
</tr>
<tr>
<td>Figure 11.</td>
<td>Using Grouped Displays to Examine the VWM Capacity Debate</td>
<td>84</td>
</tr>
</tbody>
</table>
Chapter I: General Introduction

Our visual world consists of a wealth of information composed of a variety of physical properties. Through important developments in vision science, the scientific community has come to understand that these physical properties require elaborate processing prior to phenomenological perceptual experience. Yet, conscious reports of the visual world do not, under ordinary instances of human perception, include piecemeal descriptions of physical stimulus properties and image statistics. Rather, phenomenological reports often describe a rich, unified visual percept. In reality, the visual system first must detect and convert information from the physical environment into a neural signal, ultimately constructing this unified percept.

Disparate elements from the visual field, once detected, stimulating the appropriate photoreceptors of the retina. Upon the transduction of physical matter to neural signals, the cortical regions comprising the visual system carry out the processing of visual information. Retinotopic organization within early visual areas of the brain allows for the creation of a patchwork of disparate elements. The detection of low-level image properties gives rise to the construction and extraction of a figure from its background. A discernable object often emerges from the combination of inputs ultimately creating the potential for entry into conscious awareness. Finally, attention can be directed to relevant visual stimuli and, when the stimuli are no longer present in the visual field, representations can be formed and actively maintained in working memory for a brief period of time.
However, prior to the allocation of attentional resources to relevant aspects within the visual field, perceptual organization within stimulus arrays serves to link otherwise disparate elements together. Gestalt principles of grouping have long been known to facilitate perceptual processing of visual information. Recently, experiments have indicated that Gestalt principles can facilitate visual working memory (VWM) performance as well. The underlying neural mechanisms and processing stages at which various Gestalt grouping principles may provide benefits to VWM, however, have yet to be fully characterized.

This dissertation will examine the role of Gestalt grouping principles in VWM processes using both behavioral methods and electrophysiological techniques. The following sections will provide the reader with an overview of working memory in general, capacity limitations, current debates within the VWM literature, neural estimates of VWM capacity, and finally will review known benefits of Gestalt grouping principles for visual perception, attention, and working memory. After providing a review of the relevant literature from these areas of research, a series of experiments will be proposed to examine the impact of Gestalt grouping principles on active VWM representations.

Working Memory

Our ability to form and temporarily maintain internal representations of our external world relies on the cognitive process known as working memory. William James (1890) first distinguished primary memory, limited with respect to capacity, from a theoretically unlimited secondary memory. Atkinson and Shiffrin (1968) later expanded upon this initial distinction by including the terms sensory, short-term, and long-term
memory in their Multi-Store Model. Today these concepts have evolved into what cognitive psychologists and cognitive neuroscientists refer to as short-term memory (STM) or working memory (WM) and long-term memory (LTM). A distinction between WM and STM is often made in the literature in an attempt to dissociate active versus passive storage mechanisms, respectively (e.g., Baddeley, 1986; Baddeley & Hitch, 1974). Recent researchers consider these terms largely interchangeable given that both constructs effectively measure the same stages of the memory process (see Unsworth & Engle, 2007b, for a review). In the forthcoming content this cognitive process will be referred to as working memory (WM).

The predominant model of WM for the past ~40 years has been the Multiple-Component Model, proposed by Baddeley and colleagues (e.g., Baddeley, 1986; Baddeley & Hitch, 1974; Baddeley & Logie, 1999). The Multiple-Component Model suggests that distinct capacity-limited systems within WM exist to process and maintain information encoded via various sensory modalities.

In contrast to the Multiple-Component Model, Cowan and colleagues proposed an alternative model of WM known as the Embedded-Processes Model (Cowan, 1999). The Embedded-Processes Model considers items held in what has traditionally been referred to as WM, to be activated LTM representations. Furthermore, this model assumes that relevant pieces of stored information within LTM can be simultaneously activated for a brief period of time. Finally, similar to the Multiple-Component Model, the Embedded-Processes Model proposes that a limited amount of temporarily activated information within LTM can be brought into the focus of attention for further processing by a central executive system (Cowan, 1999). Indeed, these two perspectives are not necessarily
mutually exclusive. For example, Cowan (2001) considers the focus of attention to be the limiting factor inherent to WM capacity. However, Baddeley’s (1986) central executive ultimately determines what information enters the focus of attention via the allocation of available resources based on task demands. Regardless of the origin of the capacity limitations within WM, examining the characteristics of these limitations is a necessary step toward developing a comprehensive understanding of this important cognitive process (Baddeley, 1986; Cowan, 2001).

**Visual Working Memory Capacity Limitations**

Within the VWM literature, a large amount of research has focused on the stages associated with the processing of visual information detected from the environment. Upon detection of stimuli in our immediate environment, visual information must be encoded for further processing beyond that which is available from basic sensory or iconic memory (> 500 ms: Averbach & Sperling, 1961; Phillips, 1974; Phillips & Christie, 1977; Sperling, 1960). Beyond iconic memory, once information is encoded into VWM, internal representations of external visual information are formed and actively maintained for a brief period of time (on the order of seconds). Finally, during a comparison process, the information being maintained can be evaluated against other external visual content or existing internal representations.

In the process of extracting information from our visual world, we are constantly prone to brief interruptions in the form of rapid eye movements known as saccades. In order to complete everyday tasks, the human visual system must sample the visual environment in an active fashion, resulting in the generation of an average of ~3 saccades
per second (Buswell, 1935; for a recent reviews, see Henderson & Hollingworth, 1998; Ross, Morrone, Goldberg, & Burr, 2001). One important issue for vision science, given these frequent eye movements, relates to the isolation of the underlying mechanisms associated with transsaccadic integration. Evidence from the literature suggests that the processes underlying the maintenance and ultimate integration of visual information across saccades are, perhaps, tantamount to visual working memory processes (Irwin, 1991, 1992). How do we construct an overall representation of the current visual scene across momentary fixations accomplished between saccades? Conceivably, the information extracted from the current visual scene during a given fixation is stitched together with information garnered between saccades at subsequent fixations (Breitmeyer, Kropfl, & Julesz, 1982). Despite this view, in tasks requiring the overt detection of changes to the visual scene occurring during a saccade, observer performance is impoverished (Henderson, 1997). On the other hand, measures of covert detection (e.g., fixation duration) of targets across saccades suggest that information from a given visual scene is represented in a highly detailed manner (Henderson & Hollingworth, 2003).

Fortunately, VWM aids in the temporary storage of information extracted from the current scene across these interruptions. Yet, despite the prevalence and importance of this cognitive process in nearly all visual tasks carried out in our daily lives, VWM is a limited-capacity process. Initial capacity estimates were first made famous by George Miller (1956) in his seminal review, “The magical number seven, plus or minus two: Some limits on our capacity for processing information.” This paper not only revealed that humans are limited in their ability to process large amounts of information, but that it is possible to group or “chunk” semantically related information into more manageable
units. Miller’s (1956) capacity estimate of 7 ± 2 items, though perhaps only meant to serve as a theoretical guide, has since been updated.

More recent estimates of VWM capacity converge on a common estimate, extracted using a variety of task paradigms, of a capacity limit of approximately 4 items (see Cowan, 2001, for a review). The large number of studies finding evidence in favor of a “pure” capacity limit of ~4 items can be considered to follow one or more of Cowan’s (2001) criteria. For example, some studies examine capacity limits by using overwhelmingly large stimulus arrays wherein each item can be considered a single to-be-remembered unit (or “chunk”). Examples of this approach are VWM tasks adapted (e.g., using longer stimulus presentation durations) from of Sperling’s (1960) classic whole report task. When up to ~3-4 items (e.g., colored squares) are presented, no significant decrease in overall performance is evident. However, at set sizes of > 4 items, accuracy drops significantly (e.g., Luck & Vogel, 1997).

Another common approach used to examine pure estimates of VWM capacity has been to prevent verbal recoding of the stimulus or limit rehearsal using an articulatory suppression task (e.g., repeat the word “toy” throughout experimental trials) or a running memory span (e.g., long lists are presented and the time of testing is uncertain) task (Pollack, Johnson, & Knaff, 1959; Waugh & Norman, 1965). Again, when these constraints are in place, VWM capacity is estimated to be limited to ~4 items. Other “pure” approaches to estimate capacity involve the use of multiple object tracking tasks. In these multiple-object tracking studies participants can accurately track ~3-4 items amongst a large number of distractor items (Pylyshyn et al., 1994).
A variety of empirical approaches to examining VWM capacity estimates point to a common limit of ~4 items. However, it is important to note that studies converging on a capacity estimate limited to ~4 items are considered “pure” measures of capacity limits. As such, Miller’s (1956) original capacity estimate may still be considered appropriate if applied to situations in which other sources of information (e.g., semantic information stored in long term memory), that are not capacity limited can be used to aid in the processing of information in conjunction with capacity-limited processes.

The Unit of Storage Debate

Estimates of capacity limited to the concurrent storage of ~4 items have spawned an ongoing debate regarding the characterization of the “unit of storage” within VWM. Do individual features, locations, or integrated objects comprise the unit of storage subject to capacity limitations within VWM? In their now seminal paper, Luck and Vogel (1997) found evidence supporting a “strong-object” hypothesis, indicating that integrated objects, not individual features, are the unit of storage within VWM. Specifically, in a series of experiments their group found nearly identical performance in change detection tasks requiring the storage of color-color conjunctions compared to single features (Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001). These results imply that when bound to the same object, additional features can be processed at no extra cost to VWM performance. Importantly, results from Vogel et al. (2001) showed that even though the same amount of information is present in conjunction (e.g., red square framed within a larger blue square) and disjoined arrays (e.g., blue frame & red square), when integrated within a single object at a single location in space, benefits to VWM become apparent.
The same information, when integrated within an object, optimizes VWM processes to facilitate performance.

While Luck and Vogel (1997) found evidence supporting the “strong-object” perspective, recent evidence suggests that modifications to this perspective are necessary. Several independent groups have failed to replicate the object benefit for color-color conjunctions (Olson & Jiang, 2002; Parra, Cubelli, Della Sala, 2011; Wheeler & Treisman, 2002; Xu, 2002b). For example, these researchers reported that VWM accuracy for color-color conjunctions was impoverished compared to single simple features (Olson & Jiang, 2002; Wheeler & Triesman, 2002). However, when color-orientation conjunctions were tested (e.g., a green bar oriented at a tilt of 45 degrees), performance was equivalent between conjunction conditions and simple feature conditions (Olson & Jiang, 2002; Wheeler & Treisman, 2002; Xu, 2002b). A logical explanation has been proposed to account for the lack of observance of equivalent performance for different types of conjunction stimuli and simple feature stimuli. This “multiple-resources” perspective suggests that competition for available resources occurs between features from the same stimulus dimension (Wheeler & Treisman, 2002; Wolfe et al., 1990). Thus, if separate “pools” aid in the maintenance of distinct features in VWM, then adding features from the same dimension (e.g., color) will deplete the remaining available resources within that pool. However, if separate pools of resources support the storage of features from distinct dimensions, VWM performance for integrated feature conjunctions from distinct dimensions (e.g., a green bar oriented at a tilt of 45 degrees) should be equivalent to performance for single features (Olson & Jiang, 2002; Wheeler & Treisman, 2002, Xu, 2002b).
Even when conjunction stimuli are composed of features from distinct dimensions, the strength with which these features are joined determines whether VWM performance will be equivalent to performance for single features. In change detection tasks requiring participants to monitor either one (e.g., orientation or color) or two features of conjunction stimuli (e.g., orientation and color), performance decreases when two features are monitored (Xu, 2002a). Take, for example, the Saturn-like conjunction stimuli used by Xu (2002a). In the various feature-monitoring conditions, the circular shape of the stimulus must be monitored for a potential color change or the orientation of the black bar must be monitored, or both features must be monitored. In another conjunction condition, colored bars with a given orientation must be monitored for a change to either the color, the orientation, or both. An additional disjunction condition requires participants to monitor the same number of colored circles, orientated black bars, or both. This pattern of results indicates that the benefits of objecthood for VWM performance do not occur in a robust, all-or-none fashion. Rather, the strength of the object-based benefits observed in VWM experiments is ultimately dependent upon the organization and spatial configuration of the items within the stimulus array (Jiang, Olson, & Chun, 2000; Olson & Jiang, 2002).

The Structure of Visual Working Memory Debate

Previous attempts to isolate the “unit of storage” within VWM have led to another active debate within the literature. Despite evidence that the fidelity of VWM representations is similar for integrated objects containing multiple features compared to simple features, an alternative view is that VWM capacity is ultimately constrained by a
fixed item limit. The rationale for this perspective comes from experiments employing VWM change detection tasks.

Results from these tasks show a monotonic decline in performance as the number of items requiring storage (set size) increases (Luck & Vogel, 1997; Pashler, 1988; Vogel, Woodman, & Luck, 2001). In these experiments, in addition to measurements of raw accuracy, estimates of capacity are also derived using Cowan’s (2001) formula, wherein capacity (K) is calculated given the set size (S) multiplied by the difference between the proportion of hits (H) and the proportion of false alarms (F), $[K = S \times (H - F)]$; adapted from Pashler, 1988]. Capacity estimates often asymptote at set sizes including ~4 items (Cowan, 2001; Luck & Vogel, 1997; Vogel et al., 2001). In this view, a limited number of “slots” are available for item storage and, once occupied, no additional items can be stored at the same time.

Born out of the perspective that VWM capacity is determined by a fixed item limit, discrete-resource or “slots” models propose that the overall number of items that can be simultaneously stored, not the total amount of information, best defines capacity limitations (Rouder et al., 2008; Zhang & Luck, 2008). As such, a fixed number of distinct items from a given stimulus array, regardless of complexity, are each assigned to a “slot” (Awh, Barton, & Vogel, 2007; Barton, Ester, & Awh, 2009). Once all of the available “slots” are occupied no additional information about any remaining items from the stimulus array can be encoded and stored within VWM (Zhang & Luck, 2008).

Other views of the discrete-resource perspective include the slots+resources model (Anderson et al., 2011; Zhang & Luck, 2008) and the slots+averaging model (Zhang & Luck, 2008). The slots+resources model suggests that the resources underlying
the number and resolution of item representations are independent. The slots+resources model suggests that the decline in precision as items are added, even for subspan arrays (e.g., set sizes 1-4), is due to the gradual depletion of a separate resource shared by distinct item representations stored in each “slot” (Anderson et al., 2011; Zhang & Luck, 2008). This differs from the slots+averaging model in an important way. The slots+averaging model presumes that an item is either stored with a certain level of precision or no representation of the item is formed. In this view, individuated items do not share a separate resource that determines the precision of each representation. Independent representations of the same item, however, can be formed and stored into separate slots. If recall of the stored item is required, the average of the independent item representations should aid in the precision of the response (Zhang & Luck, 2008).

In contrast, flexible-resource models propose that VWM capacity is determined by a common resource that can be allocated to store any number of items as required by current task demands (Alvarez & Cavanagh, 2004; Bays & Husain, 2008; Bays, Catalao, & Husain, 2009; Eng, Chen, & Jiang, 2005; Wilken & Ma, 2004). However, as more items are stored, the precision with which they are represented in VWM decreases as the overall amount of available resources is gradually depleted. According to this perspective, the storage of complex items requires a greater proportion of the available resources (Alvarez & Cavanagh, 2004). Upon review of the literature encompassing this debate, these two dominant perspectives may initially appear diametrically opposed. However, it is important to note that the slots+resources variation of the discrete resource model and the flexible resources models both propose that the precision of VWM representations decreases as the number of items increases.
Expanding upon the evidence from both discrete and flexible resource perspectives, recent empirical findings suggest that it is necessary to take higher order relationships between individual items into account when examining the structure of VWM representations. For example, while experiments often employ simple stimulus arrays containing individual items (e.g., colored squares), to examine VWM capacity, observers can encode and store summary statistics regarding the relationship between the items within stimulus arrays (Brady & Alvarez, 2011; Brady, Konkle, & Alvarez, 2011). For example, when reporting the size of a circle of a particular color category from a stimulus array with circles varying in size and color, observers will produce biased estimates based on the size of other circles sharing the same color (Brady & Alvarez, 2011). This evidence suggests that in addition to information regarding the individual items within the stimulus array, an overall “gist” is stored regarding the integration of elements within the array (Brady & Alvarez, 2011). As such, it is clear that further explicit examination of the hierarchical structure of VWM is necessary.

Each of the aforementioned models attempting to identify the structure of VWM has been supported by compelling evidence derived from novel behavioral, psychophysical, and computational modeling approaches. But what can we conclude from this ongoing debate? Are limitations in VWM capacity best described by a discrete-resource subject to a fixed item limit or by a flexible-resource that can be allocated and re-allocated based on bottom-up and top-down influences? To advance the current debate an in depth understanding of the neural signatures associated with the putative resources required for the storage of information within VWM is necessary. Presently, the number of experiments in the literature that have explicitly examined the underlying neural
signatures predicted by each of these VWM models is somewhat sparse. Fortunately, a wealth of neural evidence has recently examined the neural signatures associated with known VWM capacity limitations in general. In the following sections, these converging patterns of neural evidence collected from a variety of cognitive neuroscience techniques will be discussed.

**Neural Correlates of Visual Working Memory Capacity**

With advances in technology comes improvement in the precision of techniques at the disposal of cognitive neuroscientists to answer questions concerning the neural correlates underlying capacity limited VWM processes. Functional magnetic resonance imaging studies (fMRI) have found neural evidence correlated with previous behavioral capacity limitations. For instance, scanning participants while performing a delayed match-to-sample task results in the observance of a specific pattern of the blood oxygenation level-dependent (BOLD) response within areas in posterior parietal cortex (PPC, Todd & Marois, 2004; Xu & Chun, 2006).

As the number of items to be maintained (i.e., set size) increases, the BOLD response associated with delay-related activity in the intraparietal sulcus (IPS) and the intraoccipital sulcus (IOS) increases (Todd & Marois, 2004). The BOLD response emerging from the IPS/IOS increases during the delay period with set size until capacity is reached. At set sizes beyond an individual’s behavioral capacity limit the BOLD signal plateaus, indicating that no remaining neural resources are available to support storage of additional items (Todd & Marois, 2005). Converging evidence has recently been acquired using a related neuroimaging technique, functional near-infrared spectroscopy (fNIRS).
For example, in an fNIRS experiment examining delay-related activity, the change in concentration of oxy-hemoglobin (HbO) in regions such as the IPS/IOS (relative to baseline) was higher when a change detection task required simultaneous storage of 4 items compared to 2 items (Cutini et al., 2011).

Additional fMRI evidence has dissociated activation patterns between the inferior and superior regions of the IPS. Inferior IPS activations are associated with the maintenance of a limited number of objects at distinct spatial locations selected by spatial attention (Xu & Chun, 2006). Evidence from the same suite of experiments suggests that the superior IPS and areas within lateral occipital cortex (LOC) select especially relevant objects for maintenance. However, whereas the inferior IPS maintains a fixed number of objects, the concurrent amount of objects that the superior IPS can maintain at a given time depends upon the complexity of the selected objects (Xu & Chun, 2006).

**Electrophysiological Correlates of Capacity**

Evidence from human electrophysiology experiments using electroencephalogram (EEG) and event-related potential (ERP) techniques converges with the neural activation patterns associated with VWM capacity found in previous fMRI studies. Using ERP techniques, Klaver, and colleagues were the first to isolate a negative going slow wave emerging from posterior electrode sites on the scalp during the maintenance period of a VWM task. This characteristic negative slow wave was found at electrode sites contralateral to the hemifield in which stimuli were displayed (Klaver et al., 1999). More recently, in a variety of experiments, Vogel and colleagues have expanded this work by developing a paradigm to examine the neural correlates of VWM capacity during
maintenance (McCollough, Machizawa, & Vogel, 2007; Vogel & Machizawa, 2004; Vogel, McCollough, & Machizawa, 2005).

The EEG is recorded at a variety of standard scalp sites (10-20 system; Jasper, 1958). The neural activity associated with each experimental condition from electrode sites both contralateral and ipsilateral to the cued visual hemifield is then averaged to create contralateral and ipsilateral waveforms for each condition. Finally, difference waveforms are created representing the CDA by taking the difference between average contralateral and ipsilateral waveforms for each experimental condition (analysis procedure developed by Vogel & Machizawa, 2004). These results emphasize the contralateral organization of VWM. Items presented in the right visual hemifield (RVF) will elicit a greater response in the (contralateral) left hemisphere compared to the (ipsilateral) right hemisphere.

The common finding in such experiments is that at posterior scalp sites the CDA amplitude increases with set size until VWM capacity is reached (e.g., McCollough, Machizawa, & Vogel, 2007; Vogel & Machizawa, 2004; Vogel, McCollough, & Machizawa, 2005). When capacity limits are reached, the amplitude associated with the CDA asymptotes, indicating that no additional neural resources are available to support the maintenance of subsequently added stimuli. Additionally, the set size at which the CDA amplitude reaches asymptote is strongly correlated with individual capacity limits, indicating that this component is sensitive to individual differences in VWM capacity (Vogel, McCollough, & Machizawa, 2005).

Individual differences in VWM capacity are further revealed when measuring the CDA during tasks involving attentional filtering. For example, if four items are present
but only two are relevant to the task, the CDA corresponding to the maintenance of objects with this type of stimulus array is contingent upon an individual’s VWM capacity. Several studies have shown that CDA amplitudes for relevant items (e.g., set size 2) versus relevant items plus distractor items (e.g., set size 2) are nearly identical for participants with high VWM capacity (Fukuda & Vogel, 2009; Fukuda & Vogel, 2011; Spronk, Vogel, & Jonkman, 2012; Vogel, McCollough, & Machizawa, 2005). Participants with low VWM capacity, on the other hand, exhibit a different pattern of CDA amplitudes when distractor items are present in the stimulus array.

In addition to exploring the number of discrete items that individuals can store at a given time, previous research examining the CDA measured the neural signatures associated with the storage of visual phenomena ranging from basic features (e.g., color, Vogel & Machizawa, 2004; Vogel et al. 2005, orientation, McCollough et al., 2007, shape, Diamantopoulou, Poom, Klaver, & Talsma, 2011), object complexity (e.g., Gao et al., 2009; Luria, Sessa, Gotler, Jolicœur, & Dell’Acqua, 2010), feature conjunctions (e.g., Luria & Vogel, 2011a; Woodman & Vogel, 2008), feature binding (Wilson, Adamo, Barense, & Ferber, 2012), and even high-order stimulus configurations with social relevance (e.g., faces, Sessa, Luria, Gotler, Jolicœur, & Dell’Acqua, 2011). Thus, it is evident that the CDA is an extremely useful ERP component for examining the neural signatures associated with online storage of VWM representations.

Recently, other studies produced compelling evidence that the CDA tracks the storage of object identity information bound to distinct items. For example, when item contrast is manipulated (e.g., high and low contrast colored squares), a VWM behavioral performance decrement is observed for low contrast item displays relative to high
contrast item displays (Ikkai, McCollough, & Vogel, 2010). However, the CDA amplitude corresponding to the storage of low contrast items increases from 2 items to 4 items and then asymptotes, which is the same pattern that is observed when storing high contrast items (Ikkai et al., 2010). Furthermore, when comparing the storage of 4 items presented sequentially (i.e., 2 items per array) in either the same or different location (i.e., 4 items-2 locations or 4 items-4 locations), there is no difference in CDA amplitude between the same and different location conditions (Ikkai et al., 2010). Rather, the CDA amplitude only increases when additional distinct items must be stored, independent of the number of locations in which the items are presented (Ikkai et al., 2010). Converging evidence in favor of the notion that the CDA tracks object identity information indicates that the CDA amplitude is the same for 1 item with 1 color (e.g., 1 blue square) compared to 4 items of identical color (e.g., 4 blue squares) displayed in different locations (Gao et al., 2011). Thus, results from this experiment found that the amplitude of the CDA only increased significantly when the number of distinct item identities, not locations, increased (Gao et al., 2011).

Compelling evidence suggests that the CDA stores the identity information (e.g., a blue square) associated with distinct items presented within stimulus displays. Given these findings, does the amplitude of the CDA vary with regard to the complexity of distinct item identities? In the VWM behavioral literature previously discussed, evidence exists both in favor of and opposed to the proposition that VWM capacity is dependent upon item complexity (e.g., Alvarez & Cavanagh, 2004; Awh et al., 2007). In behavioral studies, however, capacity estimates are derived from measures of performance at the time of retrieval, which may be confounded by similarity between the probed item and
the stored items during the comparison process (Awh et al., 2007). Assessing VWM capacity with respect to item complexity by measuring the CDA avoids these potential confounds because estimates of capacity are obtained during the delay period prior to the comparison process (Luria et al., 2010).

In recent electrophysiological examinations of VWM capacity for both simple and complex items, a consistent pattern of results has emerged. First, complex items such as random polygon shapes require a greater amount of neural resources, as indexed by higher CDA amplitudes, compared to simple items (e.g., basic shapes, Gao et al., 2009; colored squares, Luria et al., 2010; Luria & Vogel, 2011a). Second, this is supported by behavioral evidence that accuracy is significantly higher when remembering basic shapes compared to random polygons (Gao et al., 2009; Luria et al., 2010). Importantly, for simple stimuli, the CDA amplitude increased as set size increased (e.g., from 2 to 4 items) until capacity was reached. Random polygon stimuli, on the other hand, produced CDA amplitudes in which there was no significant difference between set sizes with 2 compared to 4 random polygons (Gao et al., 2009; Luria et al., 2010). Importantly, in a direct comparison of the proportion of neural resources required for the storage of 2 simple shapes compared to 2 random polygons, the CDA amplitude was found to be significantly larger when storing 2 random polygons compared to 2 colored squares (Luria et al., 2010). These electrophysiological results support previous assertions that VWM capacity limitations are set by not only the number of items, but also by the complexity of the information being stored (Alvarez & Cavanagh, 2004).

The validity of the CDA as an index specific to the storage of VWM representations rather than as an index of simpler, low-level electrophysiological events,
such as eye movements, has recently been called into question. The CDA is measured during a task requiring participants to maintain central fixation throughout the task. Perhaps the most difficult aspect of these unilateral change detection tasks is maintaining fixation while an endogenous cue indicating the relevant side of the memory array is presented (e.g., < left or right >). Even when participants comply with the task, their eyes may deviate, involuntarily, from a fixation cross (e.g., within 0.2° to 2.0°) due to small eye movements (e.g., microsaccades; see Martinez-Conde, Macknik, Troncoso, & Dyar, 2006; Rolfs, 2009). Artifacts due to microsaccades are difficult to remove from EEG/ERP data and can influence brain potentials recorded from posterior sites along the scalp (Dimigen, Valsecchi, Sommer, & Kliegl, 2009; Yuval-Greenberg, Tomer, Keren, Nelken, & Deouell, 2008).

Given the difficulty in controlling these eye movements due to microsaccades, it is possible that the ERPs examined in CDA studies emerge from a corneoretinal potential, rather than from higher-order posterior brain regions. Indeed, recent discoveries indicate that participants move their eyes to the locations of items from a memory array that are no longer in view during the delay-period of VWM tasks (Williams, Pouget, Boucher, & Woodman, 2013). Given that the corneoretinal potential is associated with a dipole that is positive toward the cornea and negative toward the retina, it is difficult (due to the inverse problem) to rule out the possibility that the CDA is merely an artifact of eye movements (Plöchl, Ossandon, & Konig, 2012). As such, continuous gaze shifting from fixation to the cued hemifield containing items from the memory array could produce the CDA, as this ERP is observed and measured as a sustained, slow negativity over posterior scalp sites.
Recent evidence, however, suggests that this alternative “corneoretinal-potential” hypothesis, which suggests that the CDA is the result of low-level eye movement artifacts, fails to account for common findings from the ERP-related VWM literature. In an important examination of this alternative hypothesis, Kang & Woodman (2014) measured, analyzed, and plotted the time course of eye movements, the horizontal electrooculogram (HEOG), and the CDA during a VWM color change detection task that varied with respect to set size. They found that the amplitude of the CDA increased as a function of set size, but did not continue to increase at set sizes beyond behaviorally estimated capacity limits (e.g., > ~3 items). Importantly, both the amplitude of the HEOG and the degree of deviation of eye gaze from fixation continued to increase at set sizes beyond VWM capacity limits (e.g., increased from 4 to 6 items). As such, involuntary eye movements (e.g., microsaccades) in the direction of stored item locations can be elicited during the delay-period of VWM tasks. However, such eye movement artifacts are \textit{byproducts} of spatial representations being stored in VWM rather than \textit{generators} of the CDA (Kang & Woodman, 2014).

As discussed above, the number and the complexity of items within a memory array alter the CDA. Yet, despite this wealth of electrophysiological evidence another fundamental, but important, question regarding this prevalent and beneficial ERP component remains unanswered. Does the CDA reflect the measurement of a neural resource responsible for the storage of a fixed number of items dependent upon complexity? Or might the CDA also measure a neural resource that is flexibly allocated based on the perceptual organization inherent to a given memory array? One possibility is that basic principles of perceptual organization prevalent in stimulus arrays may, for
better or worse, alter the CDA. For example, if the presence of Gestalt principles in stimulus arrays were beneficial for VWM processes, perhaps fewer neural resources would be required to store grouped stimuli. This intriguing possibility would also indicate that the presence of perceptual grouping cues could result in VWM performance improvement compared to ungrouped arrays. In the following sections, Gestalt principles of perceptual grouping and their influence on perception and VWM processes are discussed. Finally, a series of experiments were designed and conducted in an attempt to further explore the neural signatures associated with the storage of grouped item representations within VWM.

**Gestalt Principles of Grouping**

Psychological researchers have long been fascinated in the process by which the human visual system constructs a rich, integrated percept from an initially chaotic collection of photons at the retinal level. In particular, Max Wertheimer, along with colleagues Kurt Koffka and Wolfgang Köhler, founded Gestalt psychology based on their interest in the principles of perceptual organization. Gestalt psychologists developed the theory that our phenomenological experience is characterized by perceiving the world according to a variety of principles that serve to group distinct elements into an integrated percept. Wertheimer (1924/1950) first demonstrated a variety of these Gestalt principles of grouping by altering the configuration of basic stimulus displays (e.g., circles). When identical stimuli are equally spaced, no grouping cues are available to aid in the organization of the display; see Figure 1a (Wertheimer, 1924/1950). When stimulus displays are organized such that Gestalt principles of grouping are present, however,
grouped elements appear to “belong together” (Rock, 1986). If the same stimuli are arranged such that some are physically closer while others are farther apart, the principle of proximity becomes apparent, serving to group the individual items; see Figure 1b (Wertheimer, 1924/1950). Repetition of a given stimulus feature (e.g., color, shape, orientation, size) results in the grouping of equally spaced items, highlighting the principle of similarity; see Figure 1c (Wertheimer, 1924/1950). Later discoveries highlighted the importance of uniform connectedness, which can override the principles of proximity and similarity by physically linking items within a stimulus display into a single object; see Figure 1d (Palmer & Rock, 1994). The impact of these three Gestalt principles on VWM processes will be the focus of this dissertation.

Other classic Gestalt principles include common fate, good continuation, and closure (Wertheimer, 1924/1950). The first of these extends the range of Gestalt principles from static to dynamic stimulus displays. Referred to as common fate, stimuli moving at the same speed in the same direction appear grouped; see Figure 1e (Wertheimer, 1924/1950). Line segments that appear to continue as a single line despite being intersected with other line segments characterize the principle of good continuation; see Figure 1f (Wertheimer, 1924/1950). Finally, the principle of closure can override good continuation in some instances via the physical completion of line segments within the stimulus display; see Figure 1g (Wertheimer, 1924/1950). These grouping principles tend to ease the perceptual processing of not only contrived stimulus displays, but also aid in constructing our phenomenological experience of the world.
Figure 1: An Illustration of Gestalt Principles of Grouping

![Illustration of Gestalt Principles of Grouping](image)

*Figure 1 caption*: Illustrations of prevalent principles of perceptual grouping: A) ungrouped items, B) grouping via proximity, C) grouping via similarity, D) grouping via uniform connectedness, E) grouping via common fate, F) grouping via good continuation, G) grouping via closure.

**Gestalt Principles in Visual Perception Research**

Many experiments within the literature have characterized the manner in which Gestalt principles of grouping aid in visual perception. First, segmentation of the visual field into distinct objects occurs based on the presence of Gestalt grouping principles (Duncan, 1984; Kahneman & Henik, 1977; Neisser, 1967). Importantly, the segmentation process facilitated by the presence of Gestalt principles is thought to occur preattentively (Duncan, 1984; Duncan & Humphreys, 1989; Kahneman & Treisman, 1984; Moore & Egeth, 1997; Neisser, 1967). It should be noted that several findings in opposition to this perspective have shown that, in the absence of attention, rather obvious grouping patterns
often go unnoticed (Ben-av, Sagi, & Braun, 1992; Mack & Rock, 1998; Mack, Tang, Tuma, Kahn, & Rock, 1992). For instance, in variations of the inattention paradigm, observers are charged with a demanding discrimination task of determining which line of a briefly presented central fixation cross was longer (e.g., horizontal or vertical, Mack & Rock, 1998; Mack et al., 1992). On some of the trials, unexpected objects or grouped elements will also appear in addition to the central fixation cross. When asked to recognize the orientation (e.g., horizontal or vertical) or mere presence of a grouped array after making a length discrimination regarding the central fixation task, observers often fail to accurately recognize the grouped elements presented during inattention trials (Ben-av et al., 1992; Mack et al., 1992).

In support of the notion that grouping occurs in the absence of attention, results from experiments using variants of the inattention paradigm have demonstrated that observers can make implicit perceptual judgments based on the presence of grouping cues (Driver, Davis, Russell, Turatto, & Freeman, 2001; Lamy, Segal, & Ruderman, 2006; Moore & Egeth, 1997; Russell & Driver, 2005). In an innovative series of experiments, Moore and Egeth (1997) presented participants with two black horizontal lines surrounded by a matrix composed of mostly white dots and a few black dots. Participants were required to indicate which horizontal line was longer. During some of the trials, the black dots were randomly arranged within the matrix. In other trials, however, the black dots, grouped via similarity, together with the two horizontal black lines formed either the Ponzo illusion (e.g., line nearest converging end of two lines appears longer) or the Müller-Lyer illusion (e.g., line connected to arrowheads facing inward appears longer). In inattention trials in which the dots were grouped via similarity
to form either of these illusions, line length discriminations, driven by the presence of the illusions, were more accurate than would be expected by chance (Moore & Egeth, 1997). Thus, although participants could not explicitly identify the dot patterns when asked to do so, the presence of similarity influenced perceptual discriminations implicitly. Using similar paradigms, others have replicated and extended these important findings by demonstrating the occurrence of Gestalt principles of grouping in the absence of attention (e.g., similarity of color & uniform connectedness, Lamy et al., 2006; similarity of color, Russell & Driver, 2005).

While the presence of grouping cues facilitates the processing of visual arrays, all Gestalt principles are not equally effective. For instance, proximity has been shown to ease the process of making visual discriminations to a greater extent than similarity (Ben-Av & Sagi, 1995; Han, 2004; Han, Humphreys, & Chen, 1999; Quinlan & Wilton, 1998). Additionally, proximity has been shown to occur prior to similarity (e.g., Ben-av & Sagi, 1995). When horizontal or vertical line discriminations must be made on the presence of conflicting similarity and proximity cues, proximity dominates when the stimuli appear for a very brief (e.g., <100 ms) duration (Ben-av & Sagi, 1995).

In other work, the benefits of grouping for perception only become apparent when multiple principles are combined within a single array. For example, when global letters (e.g., H versus E) composed of local grouped elements (e.g., small circles creating an H within a matrix of squares) must be discriminated; similarity of shape alone was not sufficient to facilitate performance (Han et al., 1999). When the local elements are not only similar in shape but also grouped by uniform connectedness, performance benefits are evident in the form of faster reaction times relative to similarity alone (Han et al.,
Proximity, on the other hand, produced reaction time benefits regardless of whether it was paired with uniform connectedness. Other work, however, has shown that proximity and similarity together can produce an additive benefit in perceptual discrimination tasks, such that the benefit of similarity and proximity combined is twice that of each principle when isolated (Kubovy & van den Berg, 2008). As such, it is clear that the effectiveness of each Gestalt principle to benefit visual perception varies.

**Gestalt Principles in Visual Working Memory Research**

Gestalt principles of grouping have been shown to enhance perceptual processing by segmenting the visual field and subsequently improving measures of perceptual performance (e.g., discrimination accuracy, reaction time, Ben-av & Sagi, 1995; Chen, 1986; Han et al., 1999; Quinlan & Wilton, 1998). In many instances of everyday perception, relevant visual stimuli briefly appear and then disappear. It is often the case that knowledge of the details inherent to these visual stimuli is essential to accomplish everyday tasks. Fortunately, active representations are formed and maintained within VWM in the absence of the visual stimulus and despite momentary interruptions from eye movements (e.g., macro- and micro-saccades) and blinks (see Hollingworth, Richard, & Luck, 2008). First, a visual stimulus must be perceived (i.e., during encoding) prior to being actively stored within VWM. Given that Gestalt principles facilitate perceptual processing and subsequent measures of performance, it seems logical that these grouping cues could enhance VWM processes.

Indeed, several recent studies within the VWM literature have shown that Gestalt principles of grouping improve behavioral measures of VWM performance (Peterson &
Berryhill, 2013; Brady & Tenenbaum, 2013; Woodman, Vecera, & Luck, 2003; Xu, 2002, 2006; Xu & Chun, 2007). Among these studies, uniform connectedness and proximity have each been used in separate experiments in conjunction with a cuing paradigm to bias encoding toward grouped items within the stimulus array (Woodman et al., 2003). As such, when an explicit cue is used to bias encoding toward the location of a subsequently grouped item, accuracy improved significantly when the remaining item of the grouped pair, compared an ungrouped item, was probed (6% & 12% improvement for arrays grouped via connectedness & proximity, respectively, over ungrouped arrays containing the same number of items).

The influence of connectedness and proximity on VWM performance has also been examined. For example, parametrically varying the connectedness and proximity between two features of an object (e.g., mushroom stimuli composed of circular caps of various colors and stems of various orientations) created the possibility of examining accuracy for trials requiring monitoring a change to either a single feature or both features (Xu, 2002, 2006). An additional influence of uniform connectedness on VWM processes relates to items sharing a common region. VWM performance was higher for items (e.g., shapes) initially presented within the same common region compared to items in distinct regions (Xu & Chun, 2007).

In addition to previous findings that proximity, uniform connectedness, and common region are all beneficial to VWM performance, another prevalent grouping cue, similarity, has recently been examined. VWM accuracy is higher for stimulus arrays containing two items (e.g., circles) sharing the same color compared to arrays with the same number of items but no color repetition (Peterson & Berryhill, 2013). This grouping
benefit of similarity was consistent across set size (e.g., 3, 4, & 6 items). Additionally, observers were more confident in their responses during a change detection task for grouped trials compared to ungrouped trials. Importantly, the benefit of similarity for VWM was constrained by the principle of proximity. Thus, the presence of proximal color similarity improved VWM performance (9% improvement), but intervening items eliminated the benefit (Peterson & Berryhill, 2013).

Surprisingly, recent neural evidence has shown that the presence of similarity (e.g., orientation and color) does not alter the CDA for grouped compared to ungrouped stimulus arrays (Shen, Yu, Xu, & Gao, 2013). Both a three-item-two-color condition and a two-item-two-color condition led to greater VWM accuracy compared to a three-item-three-color condition. However, while the CDA amplitude was higher for three compared to two items, there was no difference in amplitude between the grouped and the ungrouped three-item conditions. Thus, despite behavioral evidence showing benefits to VWM performance via similarity of orientation and color, the ERP results in this series of experiments indicated no reduction in the CDA for grouped compared to ungrouped arrays of the same set size (Shen et al., 2013).

The finding that grouping via similarity enhances behavioral performance but has no effect on the CDA is surprising. Indeed, if a set size of three objects and two colors produces identical behavioral performance to that of two distinctly colored items, one would expect the amount of neural resources required to store the grouped three-items and ungrouped two-items to be similar. One possible explanation for this lack of a difference in CDA amplitude relates to the influence of incidental differences in the strength of proximity and similarity present in the grouped arrays. For example, recent
evidence suggests that similarity benefits VWM performance only when the two grouped items sharing the same color are not separated by spatially intervening, distinctly colored items (Peterson & Berryhill, 2013). This evidence suggests that the spatial proximity between items sharing the same color constrains the effectiveness of grouping via similarity.

Often, studies that employ change detection tasks to examine VWM performance unknowingly introduce proximity and similarity of color into stimulus arrays by randomly choosing from a set of colors with replacement (e.g., Luck & Vogel, 1997). As such, any stimulus array including two or more items (e.g., colored squares) may produce repetition of color. If the location of these items is also randomized, from trial to trial, the presence of similarity and proximity within a given stimulus array is highly variable.

Given the results previously discussed, the strength of similarity and proximity present in stimulus arrays is an important factor to consider. In a recent VWM change detection experiment, researchers monitored the degree of repetition present in the colors assigned to eight squares chosen randomly with replacement (Brady & Tenenbaum, 2013). Item locations were randomly placed within an invisible 5 X 4 grid. When analyzing accuracy relative to the strength of proximity and similarity of color present in the various arrays, this study found that the strength of these grouping cues was proportional to VWM performance. Thus, the arrays in which no grouping of the items via similarity or proximity was available (e.g., no color repetition) were the most difficult, as indicated by lower accuracy for these arrays (Brady & Tenenbaum, 2013).

Systematic control over the strength of similarity and proximity cues present in stimulus arrays is important when assessing the influence of these Gestalt principles on
VWM performance. In their investigation of the influence of similarity of color on the amplitude of the CDA, Shen et al., (2013) did not control for the proximity between the two items sharing the same color present in their grouped condition. Thus, their three-item grouped condition containing two items sharing the same color confounds two distinct conditions. The first condition would include arrays in which the locations of the two items sharing the same color are spatially proximal. The second condition would include arrays in which the distinctly colored item intervenes, spatially separating the two items sharing the same color. Thus, it is quite possible that different CDA amplitudes are produced when storing these distinct configurations, which were confounded within a single condition in the Shen et al., (2013) experiments. If this were the case, it would further elucidate the lack of a difference in the amplitude of the CDA between the grouped and ungrouped conditions used by Shen et al., (2013), which both contained three items.

In support of this notion, existing neural evidence has shown that fewer neural resources are required to store grouped compared to ungrouped arrays containing the same number of items. In an fMRI experiment, Xu & Chun (2007) found that arrays containing items sharing a common region were associated with lower-amplitude activations in the inferior IPS during maintenance compared to the same number of ungrouped items. Similarly, Anderson, Vogel, & Awh (2013) found smaller CDA amplitudes for conditions in which the orientation of items (e.g., wrench-head stimuli) formed collinear groups, compared to items randomly oriented relative to one another.

Contrary to observations of grouping benefits to VWM, other evidence suggests that, under certain conditions, even salient Gestalt grouping cues (e.g., common fate and
proximity) fail to produce VWM benefits. For instance, a recent study presented arrays containing either single color or color-color conjunction square stimuli (used by Luck & Vogel, 1997) during VWM encoding and tracked the online neural index of the maintenance of these items over a delay-period via the CDA. Replicating previous findings (e.g., Luria & Vogel, 2011), arrays with two separate objects (e.g., single color squares) and two color-color conjunction stimuli elicited nearly identical CDA amplitudes, while the CDA amplitude corresponding to arrays containing four separate objects increased significantly (Luria & Vogel, in press).

Interestingly, even when four separate objects moved during stimulus presentation, ultimately arriving at two stationary locations to create two color-color conjunctions induced via Gestalt proximity, the amplitude of the CDA increased significantly compared to the other stimulus configurations. Importantly, only when the separate objects arrived at the same location (i.e., proximity cue) and then moved together (i.e., proximity cue and common fate cue) during stimulus presentation did the amplitudes of the CDA for the four separate objects and the two color-color conjunction conditions converge (Luria & Vogel, in press). As such, the initial representation of stimulus configurations formed during encoding is an important factor limiting the effectiveness of Gestalt grouping cues to support VWM processes.

Many of these findings suggest that the presence of perceptual grouping in visual arrays should not only produce benefits for VWM observable via behavioral performance, but should also decrease the amount of neural resources required to store grouped item representations. Observing a reduction in CDA amplitude for arrays grouped by both proximity and similarity of color relative to similarity alone would clarify the ERP results
of Shen et al., (2013). Furthermore, this pattern of results would add neural evidence in support of recent behavioral findings, which have shown that the benefits of grouping via similarity are constrained by spatial proximity (Peterson & Berryhill, 2013).

It is clear that Gestalt grouping principles facilitate perception. Additionally, recent evidence has shown that proximity, uniform connectedness, similarity, and common region are beneficial to VWM performance (Brady & Tenenbaum, 2013; Peterson & Berryhill, 2013; Woodman et al., 2003; Xu, 2002, 2006; Xu & Chun, 2007). The underlying neural correlates associated with the storage of items grouped via these principles, however, have yet to be fully explored.

**Outstanding Questions and Overview of Present Experiments**

Previous sections of this dissertation have described the literature associated with known limitations inherent to VWM capacity. Given the important functions of VWM in our daily lives, how might our ability to store information from the visual world be improved? As discussed in the sections above, select Gestalt principles of grouping facilitate VWM performance. Given that grouping cues benefit behavioral performance, does their presence alter the neural signature associated with the storage of items grouped according to Gestalt principles? If fewer neural resources are required to store grouped items within VWM, can this be measured by examining known signatures of online storage, such as the CDA? Furthermore, what is the relative strength of each distinct grouping principle toward altering the CDA?

In the following suite of experiments, we examined the neural correlates associated with VWM maintenance of stimulus arrays grouped according to several
Gestalt principles. To do so, the electroencephalogram (EEG) was recorded during a VWM change detection task and event-related potential (ERP) techniques were used to analyze and measure the CDA associated with grouped and ungrouped stimulus arrays of various set sizes. Additionally, standard measures of behavioral performance (e.g., accuracy, reaction time, estimated capacity) resulting from the VWM change detection task were assessed. In Experiment 1, a task paradigm adapted from Shen et al., (2013) was used to control for and explore the influence of spatial proximity in arrays grouped via similarity of color. Behavioral performance and the CDA associated with items grouped via the principle of similarity of color while controlling for spatial proximity was measured. In Experiment 2, the principle of uniform connectedness was examined in isolation to compare the CDA for items grouped via connectedness and ungrouped items of identical set size. Finally, Experiment 3 examined the influence of each of the principles examined in Experiment 1 and 2, namely, similarity + proximity and uniform connectedness and their impact on the CDA.
Chapter II: Experiment 1

Introduction

The goal of Experiment 1 was to examine the influence of grouping stimuli via similarity of color on the CDA while controlling for the influence of spatial proximity. Previous examination of the influence of similarity of color found no difference in CDA amplitude between grouped and ungrouped arrays (e.g., Shen et al., 2013). This previous study, however, failed to control for the influence of incidental proximity between items in their memory arrays. As discovered in previous research, spatial proximity between items sharing the same color constrains the VWM benefits of similarity (Peterson & Berryhill, 2013). If spatial proximity is indeed a factor limiting the benefits of similarity for VWM, this will be made apparent via the CDA amplitudes for each condition.

Hypotheses

Several explicit predictions were made for Experiment 1. First, it was predicted that behavioral performance would decrease with increased set size. Second, VWM accuracy is predicted to be greater for grouped compared to ungrouped arrays of the same set size. Additionally, the amplitude of the CDA was predicted to increase with set size in ungrouped conditions (e.g., 2 to 3 items). Importantly, the CDA amplitude corresponding to 3 items grouped via similarity and strong proximity (3-SSP) should be smaller than the CDA amplitude for 3 items grouped via similarity and weak proximity (3-SWP). No differences in accuracy or CDA amplitude were expected between the 3-SWP and the 3-UG conditions.
Method

Participants

Twenty-two undergraduate students from the University of Nevada, Reno participated in Experiment 1 (14 female, mean age = 22.2 years). In order to be included in group-level analyses, a minimum behavioral performance criterion of 75% response accuracy was required for each participant. As such, seven participants were excluded from subsequent group-level analyses because their average behavioral performance (e.g., proportion of correct trials < 0.75) failed to exceed this a priori criterion. All participants were right-handed, neurologically intact, and had normal or corrected-to-normal vision. Additionally, participants were given the Ishihara color plates test to assess potential deficiencies in color vision prior to beginning the experiment. All 15 of the remaining participants from Experiment 1 passed the Ishihara color plate test. The Institutional Review Board at the University of Nevada, Reno, approved all study protocols. Participants were informed of all procedures and provided written consent prior to beginning the experiment.

Apparatus

The experimental task and stimuli were created and presented with MATLAB (Mathworks, Natick, MA) using the Psychophysics Toolbox 3.0 extension (Brainard, 1997; Pelli, 1997). Stimuli were displayed on a 19-inch NEC MultiSync E1100 CRT monitor (refresh rate of 75 Hz at a resolution of 1024 X 768) via a Mac mini 2.5 GHz dual-core Intel Core i5.
Stimuli and procedure

Colored squares (0.7 X 0.7º) were randomly chosen from a set of seven colors (cyan, white, red, blue, yellow, green, magenta). Stimuli were presented within three possible locations on each side of fixation. In order to manipulate the proximity between objects, each location was held constant at an eccentricity of 5.2º from fixation. Stimuli on each side of fixation were either presented at a distance of 5.2º (strong proximity) or 9.1º (weak proximity) apart. For arrays with only two items, the onscreen locations of the two items were counterbalanced across the three possible locations on each side of fixation.

Figure 2: Experiment 1 Task Paradigm and Stimulus Configurations

Figure 2 caption: Experiment 1 task paradigm and stimulus configurations. Participants viewed a fixation cross (300 ms) and then were given an arrow cue (200 ms) indicating the side of the screen to which they should covertly attend during the trial. A variable interval (300-400 ms) preceded the memory array (100 ms) containing one of the four
conditions displayed above. The experimental conditions included stimulus displays including either two items (2-UG), three items (3-UG), three items grouped by similarity and weak proximity (3-SWP), or three items grouped by similarity and strong proximity (3-SSP). After a blank delay period (900 ms), a probed item appeared that was either the same color (“old” trials) that was originally presented or is a different color (“new” trials). Participants were given 3 seconds to respond. Note that stimuli are depicted for illustrative purposes only and do not reflect the exact dimensions and visual angles of the stimuli displayed during the actual experiment.

Each trial began with the presentation of a black fixation cross (0.4° X 0.4°, 300 ms), followed by the presentation of either a left or right facing black arrow (2.1° X 0.4°, 200 ms) above fixation to indicate the side of the memory array to covertly attend during the presentation of the memory array. After a variable delay in which only the fixation cross was visible (300 to 400 ms), a memory array composed of one of the four experimental conditions was presented (100 ms). Stimuli were presented within two rectangular areas subtending 7.1° X 12.2° of visual angle centered 4.6° to the left or right of the fixation cross on a gray background. Participants viewed the stimuli from a distance of 57 cm. Following a delay-period (900 ms) during which only the fixation point remained on the screen, a probed stimulus (and a single non-probe stimulus on the non-cued side of the screen) appeared in one of the locations on the previously cued visual field of the display until a response was made. A single probe was used to keep the current task design consistent with previous research examining the impact of grouping cues on the CDA (e.g., Gao et al., 2011; Shen et al., 2013) and to prevent participants from making their decision based solely on the overall configuration between items (e.g., see Jiang, Olson, & Chun, 2000). Participants were required to indicate whether or not a change to the color of the probed item occurred from sample to test. If no change occurred, they were asked to press the “o” key with their right ring finger on the
keyboard and if a change occurred they were asked to press the “n” key with their right index finger. On half of the trials, the color of the probed item changed (“new” trials) from sample to test, and on the remaining half, no change occurred (“old” trials). Participants were given 3 seconds to respond. If no keyboard response was registered, the trial was considered incorrect and participants were instructed to press any key to continue to the next trial.

One of the following stimulus configurations was presented during the memory array; see Figure 2. In the two ungrouped item condition (2-UG), two colored squares were presented in two of the three consistent locations on both the left and right of fixation. In the three ungrouped item condition (3-UG), three colored squares were presented both left and right of fixation. For the grouped conditions, two of the three squares shared the same color. Arrays with three items grouped via similarity and strong proximity (3-SSP) included two squares sharing the same color, which were separated by a distance of 5.2°. Finally, arrays with three items grouped via similarity and weak proximity (3-SWP) included two squares sharing the same color presented 9.1° apart.

Prior to beginning the experiment, participants completed 24 practice trials. Thirteen blocks (with the opportunity for self-paced breaks between blocks) including 48 trials per block were presented yielding a total of 624 trials with 156 trials per condition. Trial types were randomly interleaved within each block. Participants were instructed to maintain fixation during each trial and to avoid voluntary eye movements.
Behavioral Analyses

Behavioral results for all three experiments were analyzed by taking the proportion of correct trials in each condition to compare accuracy across conditions using a repeated-measures analysis of variance (ANOVA). In addition to measuring proportion correct to assess behavioral performance on the change detection tasks, reaction time was measured and analyzed using a repeated-measures ANOVA. Finally, VWM capacity, defined here as the number of discrete items that can be simultaneously stored within VWM, will be estimated and compared across experimental conditions. Capacity estimates were derived using Cowan’s (2001) formula, wherein capacity (K) is calculated given the set size (S) multiplied by the difference between the proportion of hits (H) and the proportion of false alarms (F), \[ K = S \times (H - F) \]; adapted from Pashler, 1988. Capacity estimates for each experimental condition were compared using a repeated-measures ANOVA.

Electrophysiological Recording, Processing, and Analyses

For each experiment, the electroencephalogram (EEG) was recorded throughout the duration of the VWM change detection task. The EEG was recorded at a sampling rate of 1000 Hz with a vertex (Cz) reference from 256 electrodes mounted in a HydroCel Geodesic Sensor Net (HCGSN) amplified by a Net Amps 300 amplifier and acquired using Net Station 4.5.5 software (Electrical Geodesics Inc., Eugene, OR) running on a 2.7 GHz dual-core Apple Power Mac G5. Each individual EEG dataset was filtered using finite impulse response (FIR) filters high-passed at 0.01 Hz and low-passed at 30 Hz offline using Net Station 4.5.5 software (Electrical Geodesics Inc., Eugene, OR). The EEG
data corresponding to correct trials was segmented by experimental condition using an epoch of 200 ms prior (baseline period: 200 ms) to and ending at 1000 ms after the onset of the stimulus array.

Artifact detection and rejection routines were used to identify blinks and lateral eye movements associated with electrode sites near the right outer and left outer eye. Specifically, data segments corresponding to horizontal EOG channels exceeding a threshold of >20 μV within a 200 ms sliding window (using a moving average of 80 ms starting 600 ms prior to and ending 1000 ms after stimulus onset) were excluded prior to averaging. Segments containing blinks (e.g., >140 μV within a 400 ms sliding window with a moving average of 80 ms) were excluded prior to averaging. Trials were excluded if they had residual artifacts exceeding ±75 μV from 600 ms pre-stimulus to 1000 ms post-stimulus onset. EEG data were re-referenced off-line to the average of the left (channel 94) and right mastoids (channel 190). Bad channels (e.g., >100 KΩ electrode impedance, line noise, drift) were detected and replaced using proprietary interpolation algorithms implemented by Net Station 4.5.5. software (Electrical Geodesics Inc., Eugene, OR). Finally, the segmented EEG data for correct trials from each experimental condition were then averaged to generate ERPs for each experimental condition. Baseline correction was performed using the EEG data from 200 ms prior to the onset of the stimulus array.

ERP data corresponding to posterior electrode sites of interest for each experimental condition were exported from Net Station 4.5.5 and then imported into MATLAB (version 7.12.0.39132). Additional analyses and waveform plots (e.g., difference waveforms, repeated-measures ANOVA) were generated using custom scripts
written in MATLAB. The left hemisphere standard 10-20 posterior sites that were examined included: P3, PO3, P5, P7, PO7, P9, and TP7, as well as non-standard site locations specific to the GSN 256 net (electrodes 84, 85, 95, 98, 99, 104, 105, 107). Right hemisphere standard 10-20 posterior sites included: P4, PO4, P6, P8, PO8, P10, and TP8, as well as non-standard site locations corresponding to the paired left hemisphere sites (electrodes 141, 152, 169, 171, 177, 178, 179, 189). Average waveforms for each condition at each posterior site within the electrode array of interest were generated for both left visual hemifield (LVF) and right visual hemifield (RVF) trials.

Activity from left hemisphere electrode sites during LVF trials were used to generate ipsilateral waveforms for each experimental condition. ERP data from right hemisphere electrode sites during RVF trials were used to generate ipsilateral waveforms per condition. Activity at left hemisphere electrode sites during RVF trials and activity at right hemisphere sites during LVF trials were used to generate contralateral waveforms for each condition. Finally, to generate separate CDA difference waveforms for each experimental condition at each electrode site, the ipsilateral activity was subtracted from the contralateral activity \[\text{CDA} = \text{contra} - \text{ipsi}\]. The resulting CDA amplitudes from corresponding left and right hemisphere electrode pairs (e.g., P7/P8) were averaged to create a single CDA waveform for each condition for each participant. In order to evaluate the statistical significance of the amplitude differences across experimental conditions, amplitude values from a measurement window of 400-1000 ms post-stimulus onset were extracted and averaged to create a single average CDA amplitude value for each condition for each subject. A repeated-measures analysis of variance (ANOVA) was applied to the data to determine whether the average CDA amplitudes differed by
experimental condition. Only electrode pairs in which an increase in CDA amplitude from 2 to 3 items (ungrouped conditions) was evident were included in subsequent statistical analyses.

**Behavioral Results**

To examine whether the presence of similarity and proximity helped to improve VWM performance, we analyzed several measures of behavioral response data (e.g., accuracy, estimated capacity, and reaction time) from each experimental condition. As indicated by a repeated-measures ANOVA, in Experiment 1, there was a significant difference in accuracy across experimental conditions \((F(3, 42) = 26.67, \text{MSE} = 0.022, p < 0.001, \eta^2_p = 0.66, \beta = 0.99)\). Bonferroni corrected pairwise comparisons indicated that this overall significant difference was driven by a difference in accuracy between the 2-UG and 3-UG conditions \((2-UG = 0.91, 3-UG = 0.84, p < 0.001)\), the 3-UG and 3-SWP conditions \((3-UG = 0.84, 3-SWP = 0.93, p < 0.001)\), and the 3-UG and 3-SSP conditions \((3-UG = 0.84, 3-SSP = 0.91, p < 0.001)\). There was no significant difference in the proportion of correct trials between the 3-SWP and 3-SSP conditions \((p = 0.13)\). As such, a significant grouping benefit was evident, produced by arrays with similarity + weak proximity and arrays containing similarity + strong proximity; see Figure 3a.

Given that grouping benefits in the 3-SSP and 3-SWP conditions were observed, subsequent analyses were conducted to examine accuracy as a function of probe type in trials containing grouping via similarity and proximity (i.e., the 3-SSP and 3-SWP conditions). When the probed item was previously part of a grouped pair of items during the presentation of the stimulus array, accuracy was higher than when the probed item
was previously ungrouped. A paired t-test confirmed that this difference in accuracy as a function of probe type (grouped = 0.96; ungrouped = 0.85) in the 3-SSP condition was significant ($t(14) = 6.75, p < 0.001$); see Figure 3d. The same analysis of accuracy as a function of probe type was applied to trials from the 3-SWP condition. While accuracy was, on average, higher when the probed item was previously grouped, a paired t-test revealed that the difference in accuracy for the 3-SWP probe types (grouped = 0.95; ungrouped = 0.90) was non-significant ($t(14) = 1.63, p = 0.12$); see Figure 3c.

The accuracy analyses indicated that VWM performance was improved when grouping via similarity and proximity were included in stimulus arrays. However, another important measure of performance often used in VWM experiments uses hit rate and false alarm rate to estimate the number of items being stored in VWM. To examine whether grouping increases estimates of capacity we estimated capacity for each participant at each experimental condition. There was a benefit of grouping on estimated capacity, as measured by Cowan’s (2001) K formula; see Figure 3b. A repeated-measures ANOVA revealed a significant difference in estimated capacity (K) across all experimental conditions ($F(3, 42) = 97.54, p < 0.001, \eta^2_p = 0.87, \beta = 0.99$). Pairwise comparisons (Bonferroni corrected) indicated that this overall significant difference in estimated capacity (K) was driven by a significant difference in K values between the 2-UG and 3-UG conditions (2-UG = 1.65, 3-UG = 2.10, $p < 0.001$), the 2-UG and 3-SWP conditions (3-SWP = 2.57, $p < 0.001$), the 2-UG and 3-SSP conditions (3-SSP = 2.44, $p < 0.001$), the 3-UG and 3-SWP conditions ($p < 0.001$), and the 3-UG and 3-SSP conditions ($p < 0.001$). There was no significant difference between the 3-SWP and 3-SSP conditions ($p = 0.13$) in estimated capacity. Importantly, the greatest K values were
associated with the grouped arrays, regardless of the strength of proximity between the identically colored items.

Reaction times for each condition were relatively similar, indicating no benefit of grouping on response latency (2-UG = 1,103 ms; 3-UG = 1,159 ms; 3-SWP = 1,113 ms; 3-SSP = 1,142 ms). A repeated-measures ANOVA confirmed that there was no significant difference in reaction time between conditions ($F(3, 42) = 2.23, p = 0.10$).

**Electrophysiological Results**

Complementing the benefit of grouping on behavioral performance, Experiment 1 revealed a pattern of neural evidence indicating that, compared to ungrouped arrays containing the same number of items (e.g., 3), grouped arrays required fewer neural resources during VWM maintenance. Initial plotting and observation of the CDA data from the posterior electrode sites of interest revealed three electrode pairs (TP7/TP8, P7/P8, PO7/PO8) in which an increase in CDA amplitude corresponding to the maintenance of 2 to 3 items was evident.
Figure 3: Experiment 1 Behavioral Results

Figure 3 caption: Behavioral results from Experiment 1. A-D) The abscissa depicts the conditions being compared. A) Accuracy (proportion correct) is plotted along the ordinate. B) Estimated capacity (Cowan’s K) is plotted along the ordinate. C) The two probe types for trials from the three object similarity + weak proximity condition, which could have been a colored square that was grouped or ungrouped during stimulus presentation. D) The two probe types for trials from the three object similarity + strong proximity condition, which could have been a colored square that was grouped or ungrouped during stimulus presentation. Asterisks symbolize an observed significant difference between various conditions; $\alpha = p < 0.05$. Error bars represent the standard error of the mean in each condition.
Figure 4: Experiment 1 Electrophysiological Results

![Experimental results graph]

**Figure 4 caption:** Electrophysiological results for Experiment 1. The timecourse of the CDA (in milliseconds) is depicted along the abscissa. The amplitude scale (in microvolts) of the CDA for each experimental condition is indicated along the ordinate. The timecourse of the waveform depicts the pre-stimulus baseline period (-200-0 ms) occurring prior to the onset of the stimulus array for a given trial. The stimulus array was presented for 100 ms and was followed by a blank delay interval (900 ms). The measurement window for CDA analyses included amplitude values taken from 400-1000 ms post stimulus onset. CDA amplitudes, collapsed across electrode pairs TP7/TP8, P7/P8, PO7/PO8, are depicted for each experimental condition. Mean CDA amplitudes for condition are depicted in the bar graph on the right side of the figure. Error bars represent the standard errors of the means.

A 4 X 3 repeated-measures ANOVA comparing the factors of experimental condition (2-UG, 3-UG, 3-SWP, 3-SSP) and electrode pair (TP7/TP8, P7/P8, PO7/O8) indicated a significant main effect in CDA amplitude across experimental conditions ($F (3, 42) = 5.76, MSE = 25.29, p = 0.002, \eta^2_p = 0.29, \beta = 0.93$). Bonferroni corrected pairwise comparisons revealed that the effect was driven by a significant difference in CDA amplitude between the 2-UG and 3-UG condition (2-UG = -0.63 µV, 3-UG = -1.96 µV, $p = 0.02$). Importantly, there was a significant difference in CDA amplitude between the 3-UG and 3-SSP conditions (3-UG = -1.96 µV, 3-SSP = -0.68 µV, $p = 0.04$).
Additionally, the difference in amplitude between the 3-UG and 3-SWP conditions approached significance (3-UG = -1.96 µV, 3-SWP = -0.24 µV, \( p = 0.07 \)).

There was no main effect of electrode pair (\( F(2, 28) = 0.023, \text{MSE} = 0.024, p = 0.98 \)). Moreover, there was no significant interaction between the factor of experimental condition and the factor of electrode pair (\( F(6, 84) = 1.76, \text{MSE} = 2.26, p = 0.12 \)). Given that no significant main effect of electrode or interaction between condition and electrode was present, we collapsed across the data corresponding to these three electrode pairs to yield grand averaged CDA waveforms for each experimental condition; see Figure 4.

### Table 1: Experiment 1 Means (standard deviations)

<table>
<thead>
<tr>
<th>Stimulus Array Type</th>
<th>Accuracy (Proportion Correct)</th>
<th>Estimated Capacity (K)</th>
<th>Reaction Time (Milliseconds)</th>
<th>Mean CDA Amplitude (µV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-UG</td>
<td>0.91 (0.02)</td>
<td>1.65 (0.07)</td>
<td>1,103 (95.1)</td>
<td>-0.63 (0.33)</td>
</tr>
<tr>
<td>3-UG</td>
<td>0.84 (0.02)</td>
<td>2.10 (0.12)</td>
<td>1,159 (76.8)</td>
<td>-1.96 (0.42)</td>
</tr>
<tr>
<td>3-SWP</td>
<td>0.93 (0.01)</td>
<td>2.57 (0.08)</td>
<td>1,113 (85.7)</td>
<td>-0.24 (0.30)</td>
</tr>
<tr>
<td>3-SSP</td>
<td>0.91 (0.02)</td>
<td>2.44 (0.09)</td>
<td>1,142 (94.3)</td>
<td>-0.68 (0.27)</td>
</tr>
</tbody>
</table>

**Discussion**

The results from Experiment 1 replicate and extend previous work indicating that arrays containing grouping according to the Gestalt principles of similarity and proximity, whether induced incidentally or, as in the current experiment, deliberately, can benefit VWM processes (e.g., Peterson & Berryhill, 2013; Brady & Tenenbaum, 2013).
Evidence in favor of this perspective was obtained via measures of behavioral performance and electrophysiological signatures of VWM maintenance. Behaviorally, participants were more accurate when storing grouped stimulus arrays than ungrouped arrays of the same set size (e.g., 3 items). Additionally, estimates of VWM capacity (e.g., K) were larger for grouped compared to ungrouped arrays containing the same number of items.

The electrophysiological evidence obtained indicates that these behavioral grouping benefits are accompanied by a reduction in the amplitude of the CDA during the storage of grouped arrays. Given that the amplitude of the CDA has been shown to track the number of distinct item identities, and not merely the number of spatial locations, being stored within VWM (e.g., Gao et al., 2011; Vogel & Machizawa, 2004), the current evidence suggests that the identically colored items in each grouped condition were integrated into a single representation. Moreover, the amplitude of the CDA for each grouped condition was qualitatively and quantitatively similar to the CDA corresponding to two separate objects (2-UG condition) but not three separate objects (3-UG condition).

A recent study using a similar experimental design indicated no reduction in CDA amplitude when storing items grouped by similarity (Shen et al., 2013). It is possible that they found no reduction in CDA amplitude for grouped arrays because they randomized the location of items sharing the same color, effectively reducing any consistent proximity cues between identical items. In contrast, when proximity is explicitly manipulated and the spatial locations of items within the array are held constant, behavioral VWM benefits are accompanied by a reduction in CDA amplitude. As such, the current findings may be the result of the integration of the two grouped items into a
single representation, requiring fewer neural resources during the delay-period, as was evident via the CDA.

An analysis of accuracy by probe type for trials in which one of the grouped experimental conditions (e.g., 3-SSP & 3-SWP) was presented during VWM encoding served to elucidate the observed performance benefits. Specifically, in the similarity + strong proximity condition, when the item probed at test was previously a member of the grouped pair presented at encoding, accuracy was significantly greater than when the probed item was the remaining ungrouped item. This finding converges with previous evidence indicating that the presence of Gestalt principles in stimulus arrays can bias encoding processes to favor consolidation of grouped items over ungrouped items into VWM (Peterson & Berryhill, 2013; Woodman et al., 2003).

Using the same approach to examine accuracy as a function of probe type applied to trials containing similarity + weak proximity, however, yielded a different pattern of results. Although accuracy was higher when a previously grouped compared to ungrouped item was probed at test this difference was non-significant. Given that a single, non-identical item separated the two grouped items in this condition, it is possible that, during the integration of these two identical items, VWM resources were also directed to the ungrouped item during encoding. Thus, for trials from the similarity + weak proximity condition performance was similar regardless of the previous status of the probed item (i.e., grouped or ungrouped). However, a less intriguing but plausible explanation could be used to interpret this finding. The non-identical, ungrouped item presented in stimulus arrays containing similarity + weak proximity appeared to the left or right of fixation and near the horizontal meridian. Additionally, an endogenous arrow
cue presented above fixation was used to indicate the relevant side of the stimulus array prior to stimulus onset during each trial (i.e., left or right). As such, the high proportion of correct trials corresponding to the similarity + weak proximity condition when the previously ungrouped item was probed at test could be attributed to the item appearing at a privileged location at the time of encoding.

The findings from Experiment 1 indicate that the presence of similarity and proximity within stimulus arrays can induce benefits to VWM performance complemented by a reduction in the amplitude of the CDA when storing grouped relative to ungrouped arrays containing the same number of items. However, it is possible that other static grouping principles may facilitate VWM processes by integrating grouped items into a single representation and thereby reducing the amplitude of the CDA. In Experiment 2 we created and examined the effectiveness of stimulus arrays grouped via uniform connectedness toward facilitating VWM processes.
Chapter III: Experiment 2

Introduction

In Experiment 1, we found evidence for behavioral performance benefits and a reduction in an electrophysiological signature of the number of items being simultaneously stored in VWM (i.e., the CDA) when items were grouped via similarity and proximity cues. Similarity and proximity cues often appear incidentally in experimental paradigms using VWM color change detection tasks in which item colors are selected with replacement (see Brady & Tenenbaum, 2013). But what influence do other grouping cues that are typically only explicitly implemented within experimental paradigms have on VWM performance and delay-related activity?

Previously, the presence of uniform connectedness between items within stimulus displays has been shown to benefit VWM by improving behavioral performance during change detection tasks (Woodman et al., 2003; Xu, 2002, 2006). But are recently documented behavioral performance improvements associated with the presence of uniform connectedness associated with fewer neural resources being required to store connected items? To date, the underlying neural correlates associated with these VWM benefits have yet to be examined. Experiment 1 produced evidence in favor of the notion that grouping-related VWM performance benefits are accompanied by a reduction in the CDA. As such, it seems reasonable to expect that any grouping benefits associated with uniform connectedness should be accompanied by a reduction in the CDA. In Experiment 2, we examined the influence of uniform connectedness on the CDA and measures of behavioral performance.
Hypotheses

The predictions of Experiment 2 closely follow the predictions of Experiment 1. Specifically, accuracy should decrease with increases in set size. Greater accuracy should result from three item arrays grouped via uniform connectedness compared to ungrouped three item arrays. Additionally, CDA amplitudes corresponding to the storage of two item ungrouped arrays (2-UG) should be smaller than three item ungrouped arrays (3-UG). The CDA amplitude corresponding to the storage of three item arrays (3-C) in which two of the items are physically connected should be similar to the CDA for the 2-UG condition, but should be smaller than the 3-UG condition. The key prediction of Experiment 2 is that the presence of uniform connectedness will facilitate VWM performance and decrease the amplitude of the CDA.

Method

Participants

A new group of twenty participants participated in Experiment 2 (12 female, mean age = 21.6 years). Participants were informed of the experimental procedures and provided written informed consent. Participants had normal or corrected-to-normal vision. As in Experiment 1, in order to be included in group-level analyses, a minimum behavioral performance criterion of 75% response accuracy was required for each participant. As such, five participants were excluded from subsequent group-level analyses because their average behavioral performance (e.g., proportion of correct trials < 0.75) failed to exceed this a priori criterion. All participants were right-handed, neurologically intact, and had normal or corrected-to-normal vision. Additionally,
participants were given the Ishihara color plates test to assess potential deficiencies in color vision prior to beginning the experiment. All 15 of the participants included in the group-level analyses for Experiment 2 passed the Ishihara color plate test.

**Apparatus**

The experimental task and stimuli were created and presented with MATLAB (Mathworks, Natick, MA) using the Psychophysics Toolbox 3.0 extension (Brainard, 1997; Pelli, 1997). Stimuli were displayed on a 19-inch Mitsubishi Diamond Pro CRT monitor (refresh rate of 75 Hz at a resolution of 1024 X 768) via a Mac mini 2.5GHz dual-core Intel Core i5.

**Stimuli and procedure**

The same stimuli, (0.7° X 0.7° colored squares), and procedure used in Experiment 1 were used in Experiment 2 with the following exceptions; see Figure 5. In each condition the colored squares partially occluded a black square (0.82° X 0.82”) connected to a black rectangle (0.65° X 1.8°). The width and length of the black square and rectangle was held constant across all conditions. There were 204 trials in each condition for a total of 612 trials. These stimuli were created and used in the current experiment because similar uniformly connected “wrench-like” stimuli have been used previously in object-based attention experiments and have been shown to produce a “same-object” benefit (e.g., see Watson & Kramer, 1999). In the three item connected arrays, two of the black rectangles (i.e., wrench handles) abutted one another to form a connected pair; see Figure 5. Participants in the current study referred to these stimuli as
“open-ended wrenches” (e.g., similar to the “combination wrench” stimuli used by Watson & Kramer, 1999). The remaining procedures related to the task paradigm of Experiment 2 were identical to Experiment 1.

**Behavioral Analyses**

The behavioral performance measures and analyses used in Experiment 1 were used to analyze the behavioral response data from Experiment 2.

**Electrophysiological Recording, Processing, and Analyses**

The data acquisition, data processing, criteria for artifact rejection routines, and data analysis procedures used in Experiment 1 were used to analyze the electrophysiological data from Experiment 2.

**Figure 5: Experiment 2 Task Paradigm and Stimulus Configurations**
Figure 5 caption: Experiment 2 task paradigm and stimulus configurations. Participants first viewed a fixation cross (300 ms) and then were given an arrow cue (200 ms) indicating the side of the screen to which they should covertly attend during the trial. A variable interval (300-400 ms) preceded a memory array (100 ms) containing one of the three conditions displayed above: Two items (2-UG), three items (3-UG), or three items in which two of the items were connected (3-C). After a blank delay period (900 ms), a probed item appeared that was either the same color (“old” trials) as presented originally or was a different color (“new” trials). Participants were given 3 seconds to respond. Note that stimuli are depicted for illustrative purposes only and do not reflect the exact dimensions and visual angles of the stimuli displayed during the actual experiment.

Behavioral Results

In Experiment 2, there was a significant difference in accuracy across experimental conditions ($F (2, 28) = 56.82$, $MSE = 0.051$, $p < 0.001$, $\eta^2_p = 0.80$, $\beta = 0.99$). Bonferroni corrected pairwise comparisons indicated that this overall significant difference was driven by a difference in accuracy between the 2-UG and 3-UG conditions (2-UG = 0.92, 3-UG = 0.83, $p < 0.001$). Additionally, there was a significant difference in accuracy between the 2-UG and 3-C conditions (2-UG = 0.92, 3-C = 0.82, $p < 0.001$). There was no significant difference in the proportion of correct trials between the 3-UG and 3-C conditions ($p = 1.0$). As such, no significant grouping benefit of uniform connectedness was evident; see Figure 6a.

Although no overall grouping benefit was observed, an analysis of accuracy as a function of probe type in trials containing grouping via connectedness (i.e., the 3-C condition) revealed an interesting result. Specifically, when the probed item was previously part of a connected pair of items during the presentation of the stimulus array, accuracy was higher than when the probed item was previously ungrouped. A paired $t$-test confirmed that this difference in accuracy as a function of probe type (grouped = 0.86; ungrouped = 0.77) was significant ($t (14) = 2.68$, $p = 0.02$); see Figure 6b.
Aside from accuracy, there was no benefit of grouping via connectedness on estimated capacity, as measured by Cowan’s (2001) K formula; see Figure 6c. A repeated-measures ANOVA revealed a significant difference in estimated capacity (K) across all experimental conditions \((F(2, 28) = 9.20, p = 0.001, \eta^2_p = 0.40, \beta = 0.96)\). Pairwise comparisons (Bonferroni corrected) indicated that this overall significant difference in estimated capacity (K) was driven by a significant difference in K values between the 2-UG and 3-UG conditions (2-UG = 1.69, 3-UG = 1.95, \(p = 0.003\)) and the 2-UG and 3-C conditions (3-C = 1.91, \(p = 0.03\)). There was no significant difference between the 3-UG and 3-C conditions (\(p = 1.0\)) in estimated capacity.

Reaction times were shortest for stimulus arrays with two items (2-UG = 893 ms). The reaction times corresponding to grouped and ungrouped three item stimulus arrays were similar, mirroring the accuracy data (3-UG = 975 ms; 3-C = 982 ms). A repeated-measures ANOVA confirmed that there was a significant difference in reaction time between conditions \((F(2, 28) = 28.95, p < 0.001)\). Pairwise comparisons (Bonferroni corrected) indicated that this overall difference was driven by a significant difference in reaction time between the 2-UG condition and the 3-UG condition (\(p = 0.001\)) and a significant difference between the 2-UG and 3-C conditions (\(p = 0.001\)). There was no significant difference in reaction time between the 3-UG and 3-C conditions (\(p = 1.0\)).
Figure 6: Experiment 2 Behavioral Results

Figure 6 caption: Behavioral results from Experiment 2. A) The abscissa depicts the conditions of the experiment. Accuracy is plotted along the ordinate by showing the proportion of correct trials. B) The abscissa depicts the two probe types for trials from the three object connected condition, which could have been a colored square that was grouped or ungrouped during stimulus presentation. Accuracy is plotted along the ordinate by showing the proportion of correct trials for each probe type within the three object connected condition. C) The abscissa depicts the conditions of the experiment. Estimated VWM capacity, as measured using Cowan’s K formula (2001), is plotted along the ordinate by showing the mean K values. Asterisks symbolize an observed significant difference between various conditions; $\alpha = p < 0.05$. Error bars represent the standard errors of the means.

Electrophysiological Results

The ERP results were complementary of behavioral performance, but differed in one important respect. Using the selection criteria established for Experiment 1, only electrode pairs showing an increase in CDA amplitude from 2 to 3 items (ungrouped) were included in subsequent statistical analyses. In Experiment 2, the only posterior electrode pair in which an amplitude increase from 2 to 3 items (ungrouped) was apparent was electrode pair P7/P8. As predicted, the amplitude of the CDA was smallest for the condition with two items (2-UG = -0.82 $\mu$V). Differing from our predictions, the CDA
amplitudes corresponding to the three-item ungrouped (3-UG = -1.43 μV) condition and the grouped three-item condition (3-C = -1.34 μV) were similar.

A repeated-measures ANOVA confirmed that there was an overall significant difference in CDA amplitude across the three conditions \(F(2, 28) = 3.79, \eta^2 = 0.21, \beta = 0.64\). Pairwise comparisons (Bonferroni corrected) revealed that this effect was driven by a significant difference in the CDA amplitude between the two-item and the three-item ungrouped conditions (2-UG & 3-UG, \(p = 0.001\)). Interestingly, despite the observed difference in accuracy, there was no significant difference in CDA amplitude between the two-item ungrouped condition and the three-item condition containing uniform connectedness (2-UG & 3-C, \(p = 0.22\)). Finally, there was no significant difference in CDA amplitude between the three-item conditions (3-UG & 3-C, \(p = 1.0\)).

By viewing the CDA amplitudes plotted in Figure 7, it is evident that the amplitude for three-item connected arrays is similar to the three separate item arrays during the initial portion of the delay-period, but decreases in amplitude to the level of the two separate item arrays toward the end of the delay-period. An analysis of the CDA amplitude by condition for both the early and late delay-period confirmed this apparent difference. Early in the delay-period (400-600 ms) of the task, CDA amplitudes for both three-item conditions were significantly larger than the two-item condition \(F(2, 28) = 4.51, p = 0.02, \eta^2 = 0.24, \beta = 0.72\). Specifically, the CDA amplitude for three separate items (3-UG = -1.43 μV) was higher than two separate items (2-UG = -0.80 μV, \(p = 0.01\)). Additionally, the CDA amplitude for arrays containing connectedness (3-C = -1.51 μV) was higher than for two separate items \((p = 0.04)\). During the early delay-period, there was no difference in CDA amplitude between the three-item conditions \((p = 1.0)\).
The late portion of the delay-period (800-1000 ms), however, showed a different pattern of CDA amplitude by condition \(F(2, 28) = 3.87, p = 0.03, \eta_p^2 = 0.22, \beta = 0.65\). Specifically, while the CDA amplitude for three separate items (3-UG = -1.47 µV) remained higher than for two separate items (2-UG = -0.69 µV, \(p = 0.003\)), there was no significant difference in CDA amplitude between two separate items and three items with connectedness (3-C = -1.36 µV, \(p = 0.23\)). Finally, there was no difference in CDA amplitude between the three-item conditions (\(p = 1.0\)).

**Figure 7: Experiment 2 Electrophysiological Results**

*Figure 7 caption:* Electrophysiological results for Experiment 2. The timecourse of the CDA (in milliseconds) is depicted along the abscissa. The amplitude scale (in microvolts) of the CDA for each experimental condition is indicated along the ordinate. The timecourse of the waveform depicts the pre-stimulus baseline period (-200-0 ms) occurring prior to the onset of the stimulus array for a given trial. The stimulus array is presented (100 ms, dashed line) and followed by a blank delay interval (900 ms). The
measurement window for CDA analyses included amplitude values average across the delay-period beginning at 400 ms and ending at 1000 ms post stimulus onset (dashed line).

Table 2: Experiment 2 Means (standard deviations)

<table>
<thead>
<tr>
<th>Stimulus Array Type</th>
<th>Accuracy (Proportion Correct)</th>
<th>Estimated Capacity (K)</th>
<th>Reaction Time (Milliseconds)</th>
<th>Mean CDA Amplitude (µV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-UG</td>
<td>0.92 (0.01)</td>
<td>1.69 (0.04)</td>
<td>893 (43.6)</td>
<td>-0.82 (0.16)</td>
</tr>
<tr>
<td>3-UG</td>
<td>0.83 (0.01)</td>
<td>1.95 (0.07)</td>
<td>975 (50.3)</td>
<td>-1.43 (0.18)</td>
</tr>
<tr>
<td>3-C</td>
<td>0.82 (0.01)</td>
<td>1.91 (0.09)</td>
<td>982 (53.8)</td>
<td>-1.34 (0.29)</td>
</tr>
</tbody>
</table>

Discussion

The findings from Experiment 2 illustrate an example of a failure to replicate previous observations of benefits to VWM performance for stimulus arrays grouped via uniform connectedness (e.g., Woodman et al., 2003; Xu, 2002, 2006). Contrary to our predictions, accuracy for three item arrays containing two colored squares connected via “wrench handle” stimuli was nearly identical to that of arrays containing three separate items and significantly lower than two item arrays. For trials with stimulus arrays including uniform connectedness we analyzed accuracy as a function of probe type (e.g., connected or disconnected). We found evidence of greater accuracy when the item probed at test was previously connected to another item during encoding.

Electrophysiological results revealed a significant increase in CDA amplitude when storing two compared to three separate items. Interestingly, the amplitude of the CDA corresponding to stimulus arrays including three items in which two of the items
were connected was larger compared to the two item condition, however, this difference in mean amplitude was not significant. Additionally, the difference in CDA amplitude between the three item arrays, grouped or ungrouped, was not significant.

These results suggest that the presence of uniform connectedness is not always sufficient to produce explicit grouping benefits for VWM processes. Indeed, no VWM performance benefit was evident as a result of grouping stimuli via uniform connectedness. However, our analysis of CDA amplitude as a function of delay-period interval (e.g., early vs. late), revealed an interesting pattern of results across conditions. Specifically, early in the delay-period (400-600 ms), the CDA amplitudes for the three-item conditions were nearly identical and larger than the amplitude for the two-item condition. During the late portion of the delay-period (800-1000 ms), we observed a reduction in the CDA amplitude for the three-item connected condition. While the CDA amplitude for the grouped condition was not significantly different from either the two or three ungrouped item conditions, the reduction in CDA amplitude across the retention interval is informative. This finding suggests that, contrary to our predictions, stimuli grouped via the principle of uniform connectedness may reflect partial binding of VWM representations (i.e., connected color squares).

Indeed, previous findings have shown a reduction in CDA amplitude across the delay-period for static integrated objects and dynamic objects that become bound or integrated via grouping cues. For instance, Luria and Vogel (2011) discovered that the CDA amplitude for color-color conjunction stimuli was higher than for a single color object during the early portion of the delay-period. However, during the late delay-period, they observed a reduction in the CDA such that the amplitudes corresponding to a color-
color conjunction stimulus and a single color object became nearly identical. Moreover, the CDA amplitudes for these two conditions were significantly smaller than for two separate objects containing the same amount of information as the color-color conjunction stimuli (Luria & Vogel, 2011). This evidence suggests that the process of binding item representations within VWM takes time and requires neural resources. However, once the items are bound, fewer resources are required to maintain the integrated representation. Similarly, even when distinctly colored items are integrated into the same object (e.g., color-color conjunctions) via static Gestalt proximity or dynamic common fate grouping during the course of a trial, the amplitude of the CDA persists as though separate objects are being stored in VWM. Only when these integrated objects arrive at the same location (e.g., proximity cue) and then move together in the same direction (e.g., common fate cue) is a reduction in CDA amplitude apparent (Luria & Vogel, in press). It is evident that even the presence of robust Gestalt principles within stimulus arrays can be ineffective toward the integration of grouped objects into a single representation within VWM. Supporting this recent evidence (e.g., Luria & Vogel, in press), the findings from Experiment 2 suggest that the benefits of grouping information requiring storage in VWM are somewhat elusive.

We found no benefit of uniform connectedness for VWM performance in Experiment 2. However, it is possible that, under certain circumstances, these connected stimulus arrays could improve VWM performance. Given the benefits of similarity and proximity found in Experiment 1, implementing a design including arrays grouped via similarity + proximity and arrays including connectedness between items, may provide the necessary conditions to allow connectedness to benefit VWM performance. For
instance, it may be the case that VWM performance improves via similarity and proximity because cognitive resources at the time of encoding are allocated to the grouped items. Indeed, in Experiments 1 and in previous work, we have shown that performance is elevated when an item previously grouped at the time of encoding is probed at test compared to when an ungrouped item is probed (Peterson & Berryhill, 2013). Thus, if the allocation of these cognitive resources is weighted or re-calibrated in favor of encoding grouped items within the stimulus display, connected items may also be allocated a greater proportion of these resources, leading to a VWM performance benefit as well. In Experiment 3, we explored this possibility by integrating the designs used in each of the previous experiments.
Chapter IV: Experiment 3

Introduction

It is possible that our failure to observe a VWM performance benefit in Experiment 2 was due to insufficient cognitive resources being allocated to the connected items during encoding of these stimulus displays. However, it is possible that, under certain conditions, cognitive resources at the time of encoding may be re-calibrated to allocate a greater proportion of these resources to connected items. The conditions necessary for such a re-calibration may arise from the inclusion of another type of stimulus display in which grouping cues are present. For instance, the presence of similarity and proximity cues can bias the allocation of cognitive resources to encode grouped relative to ungrouped items (e.g., Peterson & Berryhill, 2013). As such, it is possible that the inclusion of displays containing these powerful cues may be the “tipping point” needed to re-calibrate the allocation of resources toward grouped items and yield VWM performance benefits in response to stimulus displays containing uniform connectedness.

Hypotheses

Given that previous experiments, (e.g., Peterson & Berryhill, 2013; Woodman et al., 2003; Xu, 2002, 2006; Xu & Chun, 2007), which found grouping related VWM performance benefits used somewhat different paradigms, the influence of each grouping cue on VWM performance can only be loosely estimated. However, given the results observed in Experiments 1 and 2, some explicit predictions can be made with regard to expected behavioral performance and CDA amplitudes across conditions. First, given that
the results from Experiment 1 revealed similar grouping benefits for behavioral performance and CDA amplitudes for both the 3-SWP and 3-SSP conditions, the two conditions will be combined into a single condition (3-S&P) in Experiment 3. Additionally, this combined condition (referred to as 3-S&P) is predicted to produce a grouping benefit that will be evident in both the behavioral and electrophysiological data. Second, given that the results from Experiment 2 indicated that uniform connectedness did not produce a grouping benefit in terms of behavioral performance, it is possible that, even with the inclusion of the 3-S&P condition to encourage the weighting of cognitive resources toward encoding grouped items, that accuracy will still be greater in the 2-UG and 3-S&P conditions than in the 3-C condition. Third, given the ERP results from Experiment 1, we predicted that the amplitude of the CDA for the 2-UG and 3-S&P conditions would be similar due to a corresponding expected behavioral performance benefit. Given the ERP results from Experiment 2, the CDA amplitude for the 3-C condition may be somewhat larger than the CDA amplitude for either the 2-UG or 3-S&P condition. Given the results from Experiment 2, it is also possible that the CDA amplitude for the 3-C condition will initially (e.g., 400-600 ms post stimulus onset) be larger than the CDA for the 2-UG condition and decrease to the level of the 2-UG during the late delay-period (e.g., 800-1000 ms). A key benefit of conducting Experiment 3 regards the ability to compare the amplitude of the CDA for the 3-C and the 3-S&P conditions in the same experiment. Given the results of Experiments 1 and 2, it seems logical to expect that similarity together with proximity should produce CDA amplitudes that are relatively smaller than the CDA for items grouped via connectedness. On the other hand, if cognitive resources are weighted toward encoding grouped items and a
benefit of connectedness emerges, it is possible that we might observe no difference in CDA amplitude between these two conditions.

Method

Participants

A new group of eighteen participants participated in Experiment 3 (10 female, mean age = 22.2 years). Participants were informed of the experimental procedures and provided written informed consent. Participants had normal or corrected-to-normal vision. As in Experiments 1 and 2, in order to be included in group-level analyses, a minimum behavioral performance criterion of 75% response accuracy was required for each participant. As such, six participants were excluded from subsequent group-level analyses because their average behavioral performance (e.g., proportion of correct trials < 0.75) failed to exceed this a priori criterion. All participants were right-handed, neurologically intact, and had normal or corrected-to-normal vision. Additionally, participants were given the Ishihara color plates test to assess potential deficiencies in color vision prior to beginning the experiment. All 12 of the participants included in the group-level analyses for Experiment 3 passed the Ishihara color plate test.

Apparatus

The same CRT monitor and Mac mini computer with the same parameters as described in Experiment 2 were used for stimulus presentation and collection of participant responses during the VWM change detection task.
Figure 8: Experiment 3 Task Paradigm and Stimulus Configurations

Figure 8 caption: Experiment 3 task paradigm and stimulus configurations. Participants viewed a fixation cross (300 ms), followed by an arrow cue (200 ms) indicating the side of the screen to which they were to covertly attend during the trial. A variable interval (300-400 ms) preceded the memory array (100 ms) containing one of the three conditions displayed above: Two items (2-UG), three items with similarity and proximity (3-S&P), or three items in which two of the items were connected (3-C). After a blank delay period (900 ms), a probed item appeared that was either the same color (“old” trials, 50%) that was originally presented or is a different color (“new” trials, 50%). Participants were given 3 seconds to respond. Note that stimuli are depicted for illustrative purposes only and do not reflect the exact dimensions and visual angles of the stimuli displayed during the actual experiment.

Stimuli and procedure

The same color square stimuli, parameters, and task procedures from Experiment 1 and Experiment 2 were combined into a single task paradigm in Experiment 3. This resulted in the presentation of stimulus configurations used in Experiment 1 and
Experiment 2 as well as one new condition. These conditions included a 2-UG condition, a 3-C condition, and a 3-S&P condition (new); see Figure 8. In the 3-S&P condition, the proximity between the two identical colors was randomized such that on a given trial, the array could include grouping according to strong proximity + similarity or weak proximity + similarity. There were 204 trials in each condition for a total of 612 trials.

**Behavioral Analyses**

The behavioral performance measures and analyses used in Experiments 1 and 2 were used to analyze the behavioral response data from Experiment 3.

**Electrophysiological Recording, Processing, and Analyses**

The same data acquisition system, data processing procedures, criteria for artifact rejection routines, and data analysis procedures used in Experiment 1 and 2 were used to analyze the electrophysiological data from Experiment 3.

**Behavioral Results**

The behavioral results of Experiment 3 differed from the predicted pattern of results described above. Behavioral accuracy in the two-item condition was similar to the results of Experiments 1 and 2 (proportion correct 2-UG trials = 0.91). Additionally, similar to the behavioral results of Experiment 2, there was no VWM performance benefit of connectedness (proportion correct 3-C trials = 0.83). Surprisingly, diverging from our previous findings (e.g., Peterson & Berryhill, 2013) and the results of the current Experiment 1 there was no benefit of grouping via similarity and proximity on
accuracy (proportion correct 3-S&P trials = 0.74). A repeated-measures ANOVA indicated that the overall difference in accuracy between conditions was significant ($F(2, 22) = 125.48, MSE = 0.079, p < 0.001, \eta^2_p = 0.92, \beta = 0.99$). Bonferroni corrected pairwise comparisons indicated that this overall significant difference was driven by a difference in accuracy between the 2-UG and 3-S&P conditions (2-UG = 0.91, 3-S&P = 0.74, $p < 0.001$), the 2-UG and 3-C conditions (2-UG = 0.91, 3-C = 0.83, $p < 0.001$). Additionally, there was a significant difference in the proportion of correct trials between the 3-S&P and 3-C conditions ($p < 0.001$). As such, no significant grouping benefit of similarity + proximity or uniform connectedness was evident when both Gestalt principles were examined in the same experiment; see Figure 9a.

Although no overall grouping benefit of connectedness was observed, an analysis of accuracy as a function of probe type in trials containing grouping via connectedness (i.e., the 3-C condition) revealed an interesting result. Replicating the results of Experiment 2, when the probed item was previously part of a connected pair of items during the presentation of the stimulus array, accuracy was higher than when the probed item was previously ungrouped. A paired t-test confirmed that this difference in accuracy as a function of probe type (grouped = 0.86; ungrouped = 0.78) was significant ($t(11) = 2.19, p = 0.05$); see Figure 9c. The same analysis of accuracy as a function of probe type was applied to the 3-S&P trials. A paired t-test revealed that the difference in accuracy for the 3-S&P probe types (grouped = 0.73; ungrouped = 0.75) was non-significant ($t(11) = -1.80, p = 0.10$); see Figure 9d.

Aside from accuracy, there was no benefit of grouping via similarity + proximity or connectedness on estimated capacity, as measured by Cowan’s (2001) K formula; see
A repeated-measures ANOVA revealed a significant difference in estimated capacity (K) across all experimental conditions ($F(2, 22) = 32.58, p < 0.001, MSE = 0.80, \eta_p^2 = 0.75, \beta = 0.99$). Pairwise comparisons (Bonferroni corrected) indicated that this overall significant difference in estimated capacity (K) was driven by a significant difference in K values between the 2-UG and 3-S&P conditions (2-UG = 1.62, 3-S&P = 1.46, $p = 0.02$) and the 2-UG and 3-C conditions (3-C = 1.97, $p = 0.001$). Additionally, there was a significant difference in estimated capacity (K) between the 3-S&P and 3-C conditions ($p = 0.001$).

Reaction times were shortest for stimulus arrays with two items (2-UG = 825 ms), followed by the similarity + proximity arrays with three items (3-S&P = 853 ms), and, finally, the connected arrays with three items (3-C = 898 ms). A repeated-measures ANOVA confirmed that there was a significant difference in reaction time between conditions ($F(2, 22) = 29.95, p < 0.001, MSE = 0.016, \eta_p^2 = 0.73, \beta = 0.99$). Pairwise comparisons (Bonferroni corrected) indicated that this overall difference was due to a significant difference in RT between the 2-UG condition and the 3-S&P condition ($p = 0.05$), a significant difference between the 2-UG and 3-C conditions ($p = 0.001$), and a significant difference between the 3-S&P and 3-C conditions ($p = 0.004$).
Figure 9: Experiment 3 Behavioral Results

Figure 9 caption: Behavioral results from Experiment 3. A-D) The abscissa depicts the conditions being compared. A) Accuracy is plotted along the ordinate by showing the proportion of correct trials. B) Estimated VWM capacity, as measured using Cowan’s K formula (2001), is plotted along the ordinate by showing the mean K values. C) The two probe types for trials from the three object connected condition, which could have been a colored square that was grouped or ungrouped during stimulus presentation. Accuracy is plotted along the ordinate by showing the proportion of correct trials for each probe type within the three object connected condition. D) The two probe types for trials from the three object similarity + proximity condition, which could have been a colored square that was grouped or ungrouped during stimulus presentation. Accuracy is plotted along the ordinate by showing the proportion of correct trials for each probe type within the three object similarity + proximity condition. Asterisks symbolize an observed significant difference between various conditions; $\alpha = p < 0.05$. Error bars represent the standard errors of the means.
Electrophysiological Results

The ERP results from Experiment 3 were not surprising given the findings from Experiment 2 and the Experiment 3 behavioral findings. Six of the posterior electrode pairs (standard 10-20 sites: P7/P8, PO7/PO8, P9/P10; non-standard GSN 256 sites: 98/152, 107/160, 108/151) of interest showed the same overall patterns of contralateral delay activity across experimental conditions. As such, the ERP data from these six electrode pairs were averaged to yield a grand CDA for each experimental condition.

Similar to the ERP results from Experiment 1 and 2, the mean amplitude of the CDA in the two-item condition was the smallest (2-UG = -1.04 µV). The three-item condition containing grouping via similarity + proximity produced the next largest mean amplitude of the CDA (3-S&P = -1.14 µV). Finally, the largest mean amplitude of the CDA corresponded to the three-item condition including grouping via connectedness (3-C = -1.42 µV). A 3 X 6 repeated-measures ANOVA comparing the factors of experimental condition (2-UG, 3-S&P, 3-C) and electrode pair (P7/P8, PO7/PO8, P9/P10, 98/152, 107/160, 108/151) indicated no significant main effect of CDA amplitude across experimental conditions ($F(2, 22) = 1.67$, $MSE = 2.75$, $p = 0.21$, $\eta_p^2 = 0.13$, $\beta = 0.31$).

There was no main effect of electrode pair ($F(5, 55) = 1.75$, $MSE = 0.32$, $p = 0.14$). Moreover, there was no significant interaction between the factor of experimental condition and the factor of electrode pair ($F(10, 110) = 0.78$, $MSE = 0.11$, $p = 0.65$). Given that no significant main effect of electrode or interaction between condition and electrode was present, we collapsed across these six electrode pairs to yield grand averaged CDA waveforms for each experimental condition; see Figure 10.
While there was no significant main effect of condition in the electrophysiological data from Experiment 3, the results from Experiment 2 indicated differences in the amplitude of the CDA corresponding to the 2-UG and 3-C condition. As in the ERP results from Experiment 2, when viewing the CDA grand averaged waveforms (see Figure 10) from Experiment 3 it appears that the amplitude for three-item connected arrays is larger than the two-item arrays during the initial portion of the delay-period. During the late portion of the delay-period, however, the amplitude of the three connected item arrays appears to decreases in amplitude closer to the level of the two separate item arrays. As such, the lack of a main effect of CDA amplitude by condition in Experiment 3 may have been the result of averaging across potentially distinct types of maintenance processes. If this were the case, we might not be able to observe important findings regarding the influence of this grouping cue. Thus, as in Experiment 2, we conducted a CDA amplitude analysis for each condition from Experiment 3 as a function of early or late delay-period interval.

In contrast to this possibility, when analyzing the amplitude of the CDA as a function of interval within the delay period, there was no effect of early (400-600 ms) delay interval on the CDA results (2-UG early: -1.06 µV; 3-S&P early: -1.04 µV; 3-C early: -1.27 µV). A repeated-measures ANOVA indicated there was no difference in the amplitude of the CDA between any of the conditions during the early portion of the delay period \( (F(2, 22) = 0.75, MSE = 0.19, p = 0.48) \). Similarly, there was no effect of late (800-1000 ms) delay interval on the CDA results (2-UG late: -0.86 µV; 3-S&P late: -1.11 µV; 3-C late: -1.25 µV). This difference in CDA amplitude between conditions was not significant \( (F(2, 22) = 1.05, MSE = 0.46, p = 0.37) \).
Figure 10: Experiment 3 Electrophysiological Results

*Figure 10 caption:* Electrophysiological results for Experiment 3. The timecourse of the CDA (in milliseconds) is depicted along the abscissa. The amplitude scale (in microvolts) of the CDA for each experimental condition is indicated along the ordinate. The timecourse of the waveform depicts the pre-stimulus baseline period (-200-0 ms) occurring prior to the onset of the stimulus array for a given trial. The stimulus array was presented (100 ms, dashed line), followed by a blank delay interval (900 ms). The measurement window for CDA analyses included amplitude values averaged across the delay-period beginning 400 ms and ending 1000 ms post stimulus onset (dashed line).

Table 3: Experiment 3 Means (standard deviations)

<table>
<thead>
<tr>
<th>Stimulus Array Type</th>
<th>Accuracy (Proportion Correct)</th>
<th>Estimated Capacity (K)</th>
<th>Reaction Time (Milliseconds)</th>
<th>Mean CDA Amplitude (µV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-UG</td>
<td>0.91 (0.01)</td>
<td>1.62 (0.04)</td>
<td>825 (88.5)</td>
<td>-1.04 (0.17)</td>
</tr>
<tr>
<td>3-S&amp;PP</td>
<td>0.74 (0.01)</td>
<td>1.46 (0.07)</td>
<td>853 (91.8)</td>
<td>-1.14 (0.20)</td>
</tr>
<tr>
<td>3-C</td>
<td>0.83 (0.02)</td>
<td>1.97 (0.10)</td>
<td>898 (94.9)</td>
<td>-1.42 (0.23)</td>
</tr>
</tbody>
</table>
Discussion

The results from Experiment 3 indicated no benefits of grouping via similarity + proximity or uniform connectedness. Replicating the results of Experiment 2, there was no benefit to VWM performance for arrays grouped via uniform connectedness. Surprisingly, the results differed from previous findings (e.g., Peterson & Berryhill, 2013; Shen et al., 2013; Brady & Tenenbaum, 2013) and Experiment 1 with regards to the strength of grouping via similarity + proximity. Specifically, as in Experiments 1 and 2, performance in Experiment 3 was highest for the two-item arrays. Performance in the similarity + proximity condition of Experiment 3, however, was worse compared to performance in Experiment 1.

As in Experiments 1 and 2, using the Experiment 3 data we analyzed accuracy as a function of probe type in each of the two grouped conditions (3-C, 3-S&P). Replicating the results of Experiment 2, accuracy was significantly higher when the probed item was previously grouped by being connected to another item in the array at the time of encoding. In contrast to the results from the probe analysis of the data from the 3-SSP trials in Experiment 1, we did not find a difference in accuracy as a function of probe type in the similarity + proximity condition in Experiment 3.

The electrophysiological results complemented the behavioral results of Experiment 3. First, as can be observed in Figure 10, there was an increase in CDA amplitude between two-item and three-item connected arrays. Replicating the results of Experiment 2, however, this increase in the amplitude of the CDA was not significant. Second, the CDA amplitudes for the two-item condition and the three-item similarity + 
proximity condition were very similar. However, as no behavioral performance benefit of grouping via similarity + proximity was evident from the results of Experiment 3, the finding of similar CDA amplitudes for these conditions is difficult to interpret. For instance, given that behavioral performance (e.g., accuracy & estimated capacity) in the 3-S&P condition was significantly lower than in the 2-UG condition, it seems unlikely that the grouped items in the 3-S&P arrays became integrated into a single object, thereby corresponding to the observed reduction in CDA amplitude. Instead, it seems entirely possible that the observed amplitude of the CDA for the 3-S&P condition is due to an insufficient amount of neural resources being allocated to the items in the array. As such, only two distinct items from the 3-S&P array may have been stored, leading to a CDA amplitude comparable to that observed for the 2-UG condition. Given the behavioral and electrophysiological results, there was no evidence of integration of the items grouped by similarity and proximity. As such, the CDA amplitude corresponding to the 3-S&P condition likely reflects the maintenance of two distinct items.

It is surprising that simply introducing the “wrench-like” stimuli into arrays containing grouping via similarity and proximity in Experiment 3 would obliterate the robust grouping benefits observed both in the behavioral and electrophysiological results of Experiment 1. Another potential source of the disruption of grouping via similarity and proximity observed in Experiment 3 could be the orientation of the grouped items in the 3-S&P arrays. In Experiment 3, in order to hold the size of each item as well as the distance between each of the items on the left and right of fixation constant, it was necessary to rotate the orientation of the square stimuli to that of the “wrench-like” stimuli so that two of the items in the 3-C condition could be connected. As such, it is
possible that orientation may have constrained the effectiveness of the similarity and
proximity between the colored squares. However, if orientation were to completely
explain the lack of a grouping benefit for the 3-S&P arrays in Experiment 3, then this
would indicate that the benefits of grouping principles to VWM are quite fragile. At the
very least the pattern of results from Experiments 2 and 3 are interesting in light of
previous observations that grouping is automatically facilitated by the presence of Gestalt
principles in stimulus arrays (e.g., Duncan, 1984; Moore & Egeth, 1997). Instead, the
results from Experiments 2 and 3 converge with other recent CDA findings and an
emerging perspective that the initial representation of each individual item within the
stimulus array can take precedence over salient grouping cues (e.g., Luria & Vogel, in
press).
Chapter V: General Discussion

Overview of the Current Experiments

In the current experiments, Gestalt principles were examined in order to extend previous evidence finding benefits for grouped stimulus arrays requiring processing in VWM (e.g., Peterson & Berryhill, 2013; Shen et al., 2013; Brady & Tenenbaum, 2013; Woodman et al., 2003; Xu, 2002, 2006; Xu & Chun, 2007). Each of the three experiments examined VWM processes by measuring behavioral performance and delay-related neural activity during a VWM change detection task. First, the results of the current experiments provide important insights into previous findings of the beneficial outcomes of grouping items presented during a VWM change detection task according to Gestalt principles (e.g., Gao et al., 2011; Peterson & Berryhill, 2013; Shen et al., 2013; Brady & Tenenbaum, 2013; Woodman et al., 2003; Xu, 2002, 2006; Xu & Chun, 2007). Second, the current findings indicate the limitations of Gestalt principles to reflexively induce grouping between distinct items presented within to-be-maintained displays. Finally, these findings will guide future empirical efforts to examine the efficacy of traditional Gestalt principles and related bottom-up grouping cues to benefit VWM processes.

Implications of the Current Findings

In Experiment 1 we examined the behavioral and neural outcomes of grouping stimuli in VWM task displays according to the Gestalt principles of similarity and proximity. The results from Experiment 1 replicate and extend previous work showing
behavioral performance benefits of grouping via similarity and proximity (e.g., Gao et al., 2011; Peterson & Berryhill, 2013; Shen et al., 2013; Brady & Tenenbaum, 2013).

However, in contrast to previous findings (e.g., Shen et al., 2013), we found that the presence of similarity and proximity within arrays containing three colored squares (e.g., two green, one red) reduced the CDA amplitude to the level of two separate items. As such, the current findings contribute to the extant literature by indicating that VWM performance benefits are associated with a reduction the amount of neural resources required to maintain grouped stimulus arrays.

At what stage of the VWM process are these grouping benefits instantiated? While the precise time-course of these grouping benefits remains unknown, several possibilities exist. First, as in previous experiments, we found evidence of a bias to encode grouped compared to ungrouped items within VWM arrays (e.g., Peterson & Berryhill, 2013; Woodman et al., 2003). Specifically, when we analyzed measures of VWM performance as a function of which item (grouped or ungrouped) was ultimately probed at test, we found that accuracy was significantly higher when the probed item was a previously grouped, compared to ungrouped, item. This suggests that the integration of grouped items, even when presented at distinct locations, may occur at the time of encoding.

Second, in Experiment 1 we found evidence of a reduction in the amplitude of the CDA when items were grouped according to similarity and proximity. As such, the current results suggest that, when grouped according to similarity and proximity, items within VWM displays become integrated into a single representation. This converges with previous findings and suggests that the presence of Gestalt principles of grouping
may reduce the amount of neural resources required for storage within VWM (e.g., Anderson et al., 2013; Xu & Chun, 2007). For example, a recent fMRI experiment found evidence that grouped items were associated with lower amplitude responses in the BOLD signal corresponding to the IPS during maintenance when compared to ungrouped items (Xu & Chun, 2007). Similarly, supporting predictions made by discrete-resource models, a reduction in CDA amplitude is evident when stimuli form collinear groups compared to random orientations (Anderson et al., 2013). Despite this grouping advantage evident during neural signatures of VWM maintenance, there was an upper limit in the number of integrated items that could be stored, converging with previous findings in support of the discrete resource perspective (Anderson et al., 2013).

The results from Experiment 1 converge with these findings and suggest that grouped items are integrated into a single VWM representation, reducing the amount of neural resources required to maintain grouped arrays in VWM. However, previous evidence suggests that the number of integrated items that can be stored remains constrained by fixed item capacity limits (e.g., Anderson et al., 2013). As such, the findings from Experiment 1 could be interpreted as being consistent with a discrete resource perspective (e.g., Awh et al., 2007; Zhang & Luck, 2008), which is consistent with the notion that items grouped via similarity (e.g., two green squares) require the same amount of VWM resources as a one item (Anderson et al., 2013). In contrast, the current results appear inconsistent with predictions made by flexible-resource models (e.g., Alvarez & Cavanagh, 2004; Bays & Husain, 2008; Bays, Catalao, & Husain, 2009). For instance, flexible resource models seem to suggest that, if grouped items do indeed require less VWM resources, during the processing of grouped arrays, the remaining
resources could be flexibly allocated to the remaining, ungrouped items within the display. This prediction is inconsistent with our previous and current results, which found better VWM performance when a previously grouped item was probed at test compared to when an ungrouped item was probed (Peterson & Berryhill, 2013).

The grouping benefits observed in Experiment 1 could also be interpreted from a labeled Boolean map perspective (Huang & Pashler, 2007). According to this perspective, gaining conscious awareness of visual features from a given stimulus dimension is a serial process (Huang, Treisman, & Pashler, 2007). When items are grouped via similarity and share the same feature (e.g., two green squares), however, they may be mapped onto distinct spatial locations but accessed simultaneously. Accordingly, when items are grouped via similarity they would require only a single Boolean map effectively reducing the amount of resources (i.e., number of maps) required for maintenance within VWM. This perspective suggests that labeled Boolean maps may comprise the units of VWM. As such, there is an upper limit constraining the number of Boolean maps, rather than the number of discrete item representations, that can be maintained within VWM at a given time (Huang, 2010).

The current behavioral and electrophysiological findings from Experiment 1 elucidate and extend the work of recent VWM studies examining the Boolean maps perspective (e.g., Gao et al., 2011; Shen et al., 2013). For instance, Gao and colleagues (2011) found that four identically colored (e.g., four blue squares), compared to four distinctly colored, items evoked the same CDA amplitude as a single item (e.g., one blue square). This evidence suggests that the CDA tracks item identity rather than the number of distinct locations occupied by items within a VWM display. The current results are
consistent with the Boolean maps perspective indicating that VWM displays containing grouping cues via Gestalt similarity (e.g., two green squares) only require a single Boolean map and can be accessed in parallel (Huang, 2010; Huang & Pashler, 2007; Huang, Treisman, & Pashler, 2007). Finally, Experiment 1 examined set sizes presumed to be within VWM capacity limits (i.e., 3 items). However, it is likely that the ability to store arrays containing a large number of identically colored items (e.g., 6 identical color pairs) remains constrained by a fixed upper capacity limit (e.g., supporting a discrete-resource perspective). Future research will elucidate which perspective, or perhaps combination of perspectives, best accounts for the pattern of results observed in Experiment 1.

In light of recently observed grouping benefits to VWM, which of these perspectives discussed above best describes the overall capacity limitation within VWM? To examine each of these perspectives, a reasonable approach would be to follow the lead of Anderson et al. (2013) and use their recall task paradigm and “wrench head” stimuli to form collinear illusory contours to group otherwise distinct items. To illustrate the following example, when using grouped and ungrouped arrays as shown in Figure 11, each perspective would predict a different pattern of behavioral results as a function of set size. First, the flexible-resource model would predict that as more items are added, the precision with which these items can be stored decreases. As such, if this perspective best describes the capacity limitation within VWM, we would expect to see a decrease in VWM performance with increases in set size (e.g., 2 to 4 items) in the ungrouped arrays. Additionally, we might expect to observe greater precision in the grouped compared to ungrouped arrays; however, performance would still be expected to decline with
increases in set size (e.g., 2 to 4). We might also expect an increase in CDA amplitude proportionate to the amount of neural resources required to store each item within each type of display.

In contrast, the discrete-resource perspective would predict similar performance in the ungrouped conditions if an individual can hold up to 4 items within VWM at the same time. For the grouped conditions, however, if the items are grouped by either an illusory rectangle (e.g., as in set size 2) or an illusory cross (e.g., as in set size 4), we might expect to see significantly greater performance in the grouped compared to the ungrouped conditions. If the grouped items only require a single “slot” within VWM, then we would expect equivalent performance regardless of set size (e.g., 2 or 4). Additionally, when examining the CDA amplitude corresponding to the maintenance of these arrays, discrete-resource models would predict that the amplitude of the CDA would increase from set size 2 to 4 for the ungrouped arrays. However, there should be no increase in CDA amplitude with increases in set size (e.g., 2 to 4) in the grouped arrays because each group would only require a single “slot” within VWM.

Finally, given the predictions of the Boolean maps perspective, we would expect a decrease in performance from 2 to 4 items in the ungrouped arrays because these arrays require 2 and 4 Boolean maps, respectively. In contrast, this perspective would predict better performance for the grouped relative to the ungrouped arrays. Critically, performance would decrease in the grouped arrays when tasked with remembering 4 compared to 2 items. Even in grouped arrays with 4 items, two Boolean maps would be required because each illusory rectangle formed by two wrench head stimuli corresponds to a distinct feature map (e.g., vertical or horizontal orientation of the rectangle as in
Figure 11). By comparison, the grouped arrays with 2 items would only require a single Boolean map as only one illusory rectangle is formed. The CDA would increase in amplitude from grouped arrays with 2 items, followed by ungrouped arrays with 2 items and grouped arrays with 4 items, and finally would be largest for ungrouped arrays with 4 items. Such an examination using grouped and ungrouped stimulus displays might help elucidate which of these perspectives best describes the overall capacity limitation in VWM.

**Figure 11: Using Grouped Displays to Examine the VWM Capacity Debate**

*Figure 11 caption:* Examples of grouped and ungrouped stimulus configurations that could be used to test the predictions made by flexible-resource models, discrete-resource models, and labeled Boolean maps models. Each perspective proposes distinct accounts regarding the structure of the overall capacity limitation in VWM.
In contrast to the benefits of Gestalt similarity and proximity observed in Experiment 1, the findings from Experiment 2 suggest that grouping benefits to VWM via uniform connectedness are not as robust as predicted by the literature. However, given a wealth of previous evidence from the object-based attention literature, it is somewhat surprising that we observed no behavioral performance benefit of grouping via uniform connectedness. For instance, many object-based attention experiments have reported “same-object” benefits (e.g., faster or more accurate when identifying two features from the same compared to two different objects) for behavioral performance using various grouping cues (e.g., Duncan, 1984; Lavie & Driver, 1996; Vecera & Farah, 1994; Watson & Kramer, 1999). Moreover, one of these studies (e.g., Watson & Kramer, 1999) observing “same-object” benefits used “wrench-like”, uniformly connected stimuli, similar to the stimuli used in the current Experiment 2. Although the findings from Experiment 2 were somewhat surprising, based on previous evidence from the grouping and object-based VWM literature, other explanations for the current pattern of results appear plausible.

Despite observed benefits of “objecthood” obtained from the object-based attention literature, findings of object-based benefits in VWM are somewhat variable. Previous evidence indicates that, when required to temporarily hold on to multiple items in VWM, humans are just as accurate when storing two simple colored squares compared to two color-color conjunction stimuli (Luck & Vogel, 1997). However, there have been several studies that have failed to replicate these results (e.g., Olson & Jiang, 2002; Parra, Cubelli, Della Sala, 2011; Wheeler & Treisman, 2002; Xu, 2002). The common explanation for these failed attempts at replication is that when conjunction stimuli
contain two features from the same stimulus dimension (e.g., color-color), they compete for capacity and must be retrieved serially, which can result in misbinding the two features from the same dimension (e.g., Treisman & Gelade, 1980; Wheeler & Treisman, 2002).

These findings from the VWM literature provide a possible explanation for our failure to observe a grouping benefit of uniform connectedness in Experiment 2. Even though we found that, for 3-C trials, accuracy was higher when the probed item was previously grouped at the time of encoding, the overall proportion of correct trials for the 3-C and 3-UG conditions were nearly identical. Given that the stimuli used in Experiment 2 were grouped via uniform connectedness at encoding, but only a single probed “wrench” item was presented at test, it is possible that participants may have retrieved and erroneously responded to the unprobed item from the connected pair. This could account for a proportion of the response errors at the time of retrieval, negating any overall benefit originating from grouping of items at the time of encoding.

Several other possibilities exist that may explain the lack of a VWM performance benefit even when uniform connectedness was used in an attempt to group items in Experiment 2. Importantly, several differences exist between the stimuli and grouping cues used in Experiment 2 and those used in previous studies finding behavioral benefits of connectedness. For example, Woodman et al., (2003) cued the location of an item that was later connected to another item by a rectangle prior to the presentation of the stimulus display. As such, it is possible that, had we pre-cued the location of one of the connected items, we may have observed a grouping benefit. However, given other experiments finding grouping benefits associated with connectedness, explicitly directing
attention to the location of a to-be-grouped item seems unlikely to fully account for our failure to reject the null hypothesis in Experiment 2. Indeed, evidence from several experiments suggests that connecting items within a stimulus array, even in the absence of explicit direction of attention to the location of the grouped items, benefits VWM performance (Xu, 2006). One difference between the current and previous approaches relates to the influence of higher order semantic associations potentially evoked by stimuli used in previous experiments. For instance, previous evidence of the benefits of connectedness for VWM performance was obtained from a design that used two visually distinct but semantically related items (Xu, 2006). In this design mushroom cap and stem stimuli were either displayed in completely disparate spatial locations, spatially proximal to one another, or connected to one another. Performance was best when the caps and stems were connected to one another (Xu, 2006). Given the semantic association between mushroom caps and stems and prior perceptual experiences with mushroom caps and stems, perhaps such stimuli appear to “belong together” creating a more robust perceptual experience of uniform connectedness compared to the “wrench head” stimuli used in Experiment 2 (see Palmer & Rock, 1994).

Additionally, the electrophysiological results from Experiment 2 may provide an explanation as to why no grouping benefits to VWM performance were observed. First, while the CDA amplitude for 3-C arrays was larger than the amplitude for the 2-UG arrays, this difference was not significant. Second, we observed that during the early portion of the delay-period (e.g., 400-600 ms), the mean CDA amplitudes evoked by the 3-C and 3-UG arrays were equivalent and significantly larger than the CDA for 2-UG arrays. In the later portion of the delay-period (e.g., 800-1000 ms), however, a reduction
in the CDA amplitude for the 3-C arrays relative to the 3-UG arrays was observed. Thus, toward the end of the task, there was no significant difference in CDA amplitude between the 3-C and either of the remaining conditions.

This pattern of results may indicate that, even if the connected items are viewed as a grouped “unit” at the time of encoding, the distinct colors of the two connected items may never become fully integrated into a single representation during VWM maintenance. The observed reduction in the CDA for the 3-C items toward the end of the delay-period may reflect only partial integration of the connected items. Partial integration of the two connected items could result in misbinding the colored squares to the wrong location (e.g., a “wrench-head” swap) during retrieval attempts. In order to keep the current task design consistent with previous research examining neural correlates of grouping cues (e.g., Gao et al., 2011; Shen et al., 2013) and to prevent participants from making their decision based solely on the overall configuration between items we used a single probe (e.g., see Jiang, Olson, & Chun, 2000). However, given that only a single probe item was presented at test, this disruption of the original stimulus configuration within connected arrays from sample to test could have increased the probability of a misbinding error. Indeed, previous research has shown that a reduction in performance (e.g., 9%) is evident when a single probe is presented compared to when the probe is presented in the context of the original stimulus configuration (Jiang et al., 2000). The current attempt to use uniform connectedness to induce a behavioral performance benefit indicates that consistency in stimulus configuration between the sample and test period of a VWM change detection task may limit the effectiveness of this otherwise salient grouping cue.
Converging with the lack of a grouping benefit in Experiment 2, no benefit of grouping items via uniform connectedness was observed in Experiment 3. CDA amplitudes for the 3-C relative to the 2-UG condition were similar across Experiments 2 and 3. Surprisingly, in contrast to the results of Experiment 1, the presence of Gestalt similarity and proximity to induce grouping between items in Experiment 3 did not result in a VWM performance benefit. The only differences between the arrays containing similarity and proximity between Experiments 1 and 3 were the introduction of the “wrench-like” stimuli and changes to orientation of the color stimuli in order maintain consistency in size and distance between items across the conditions of Experiment 3. It is possible that slightly different orientations of the items presented in Experiment 3 may have constrained the effectiveness of the similarity and proximity between the colored squares. However, if orientation completely explains the elimination of a grouping benefit for the 3-S&P arrays in Experiment 3, then this would indicate that VWM grouping benefits are quite fragile.

At the very least the pattern of results from Experiments 2 and 3 call into question previous observations that grouping of distinct items is automatically facilitated by the presence of Gestalt principles in stimulus arrays (e.g., Duncan, 1984; Moore & Egeth, 1997). The findings from Experiments 2 and 3 support other recent evidence that the initial representation of each item in the display formed at the time of encoding plays an important role in determining whether or not the presence of salient grouping cues will benefit VWM (e.g., Luria & Vogel, in press). In the task designs of Experiments 2 and 3, the grouped items were connected at the time of encoding. At test, however, a single probe was presented, making the comparison process more difficult (e.g., converging
with the results of Jiang et al., 2000). The difficulty in retrieving the item representation from VWM that matches the probed item may arise from the disruption of the stimulus configuration from sample to test. Future work will elucidate the efficacy of grouping cues across task designs involving both the presentation of the original stimulus configuration compared to a single probe stimulus.

Limitations of the Current Experiments

Several confounds were present within the designs used in the current experiments. Some of these design flaws may help to explain the pattern of results observed in each experiment. First, in the stimulus arrays containing three items in each experiment, in order to hold the distance between each item on each side of fixation constant, one of the items was presented relatively close to the visual horizontal meridian compared to the remaining items. As such, this item could have been subject to a privileged processing status at the time of VWM encoding relative to the other items. Probing items at test that were initially presented closer to the horizontal meridian may explain, for example, the lack of an increase in accuracy as a function of probe type for the trials containing grouping via similarity and weak proximity in Experiment 1.

Second, the use of a single probe item appearing in the cued visual hemifield at test may have disrupted the initial spatial configuration of the grouped items (e.g., in the connected arrays of Experiment 2 and 3) in the array presented during encoding. Previous research suggests that the use of a single probe item, compared to a display including items presented in the same spatial configuration at test, can lead to a reduction in VWM performance (Jiang et al., 2000). We used a single probe during the test phase of the task.
for one primary reason. Previous experiments examining the neural correlates of grouping with VWM used a single probe (e.g., Gao et al., 2011; Shen et al., 2013). In an attempt to be able to compare our findings with other recent findings, we attempted, where possible, to use task procedures which were similar to these previous investigations.

Third, in Experiments 2 and 3 we used stimuli that differed from other experiments, which found that VWM performance was improved by the presence of uniform connectedness within stimulus arrays (e.g., Woodman et al., 2003; Xu, 2002, 2006). For instance, we could have used the mushroom cap and stem stimuli used in previous experiments by Xu (2006). We could have created mushroom cap stimuli of various colors attached to black stems. Moreover, in connected conditions, we could have connected the stems of two of these mushrooms during the VWM color change detection task. Perhaps the “wrench head and handle” stimuli containing the colored items in Experiments 2 and 3 were more difficult to remember due to the lack of a pre-existing association between the wrench head and handle stimuli and the colored squares. Additionally, changing the instructions at encoding to be more explicit during both the sample and test period of the task may have helped (e.g., Before the presentation of the display: “Remember the colors of the mushroom caps”. At test: “Did one of the mushroom caps change color?”). Future experiments exploring these or similar grouping principles may benefit from implementation of these factors when designing experimental task paradigms.

Finally, examining the influence of grouping cues on VWM performance concurrent with experimental designs necessary to derive lateralized ERP components
such as the CDA constrains the number of questions that can be examined within a single experiment. For instance, a sufficient amount of trials per condition (e.g., 150-200 trials) are required to produce CDA waveforms with a reasonably high signal to noise ratio, which necessarily constrains the number of conditions that can be examined in a single experimental design. Given limitations associated with EEG/ERP techniques in general (e.g., time required for electrode placement, participant wakefulness, participant motivation), the conditions that can be examined in a single experiment are limited. Additionally, in order to evoke the CDA participants must exert a great deal of effort to abstain from making saccades when presented with an endogenous cue (e.g., left or right arrow), blinking, or making excessive movements during experimental trials. Even when given regular self-paced breaks, even trained and highly motivated participants become fatigued after ~ 45 minutes. In reality, artifacts will lead to a rejection of a certain proportion of trials from each condition. To ensure that a sufficient number of artifact-free trials will be obtained for each experimental condition, it is often necessary to overestimate the number of trials that will be needed in each condition. As such, using the CDA to examine the storage of grouped stimulus arrays will be accompanied by a number of inherent limitations. Future studies with the goal of examining the influence of Gestalt principles and other grouping cues on VWM processes should take each of these factors into consideration when designing their experiments.

**Conclusions**

The discoveries from the current suite of experiments provide important contributions to the extant literature and to our understanding of the processing of
grouped stimuli within VWM. First, the results of Experiment 1 converge with previous findings from the VWM literature identifying the power of Gestalt principles (e.g., similarity & proximity) to group stimuli and improve VWM performance (Brady & Tenenbaum, 2013; Peterson & Berryhill, 2013; Woodman et al., 2003; Xu, 2002, 2006; Xu & Chun, 2007). Additionally, the results from Experiment 1 support previous findings that behavioral performance benefits are accompanied by a reduction in the amplitude of the CDA, a neural index of the number of item representations being maintained within VWM (e.g., Gao et al., 2011). Second, the current results converge with previous results showing that the automaticity of Gestalt principles to induce grouping between to-be-remembered items is constrained by several factors previously shown to be important for efficient processing within VWM (e.g., initial object representation, Luria & Vogel, in press; stimulus configuration, Jiang et al., 2000; encoding bias toward grouped items, Peterson & Berryhill, 2013; Woodman et al., 2003).

The findings from the experiments presented in the current dissertation highlight both the benefits and limitations associated with the presence of Gestalt principles within stimulus displays that require temporary storage in VWM. A variety of important visual stimuli in our immediate environment are only available for brief moments. Many of these stimuli must be actively maintained within VWM in order to complete the demands associated with both simple and complex cognitive tasks. Future VWM experiments examining these and other Gestalt principles will indicate whether imposing meaningful organization to our environment based on these principles could be beneficial to this important, but severely capacity limited, cognitive process.
References


Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia, 49*, 1632-1639.

Machizawa, M. G., Goh, C. C. W., & Driver, J. (2012). Human visual short-term memory precision can be varied at will when the number of retained items is low. *Psychological Science, 23,* 554-559.


