THE IMPACT OF STEREOSCOPIC 3-D ON VISUAL SHORT-TERM MEMORY

SARAH R. ZOHAR

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

SEPTEMBER 2015

ABSTRACT

Visual short-term memory has been studied extensively, however nearly all research on this topic has assessed two-dimensional object properties. This is unexpected, given that most individuals perceive the visual environment in three-dimensions. This thesis examines the impact of stereoscopic 3-D on visual short-term memory.

In the experiments reported here, I first investigate the stimuli necessary to directly assess visual short-term memory while eliminating two potential confounds: the use of verbal memory to encode visual information, and the unintentional use of mental resources directed at irrelevant aspects of the memory task. I found that a set of four letters sufficiently occupies verbal short-term memory; ensuring verbal memory is not used to encode visual information on the subsequent task. I also found that individuals more accurately allocate their mental resources when there is a high similarity, compared to a low similarity, between the sample stimuli being encoded and then tested. I then assess the impact of the amount of disparity, as well as the distribution of elements in depth, on visual short-term memory. Individuals retain simple visual stimuli equivalently when information is displayed in 2-D or 3-D, regardless of how objects are distributed in 3-D. Although disparity does not influence visual short-term memory, the ease of encoding does improve performance. Tasks that facilitate encoding result in better visual short-term memory performance than tasks that disrupt encoding strategies.

Taken together, the experiments reported here show that stereoscopic 3-D does not improve visual short-term memory. This finding is true regardless of how visual objects are distributed in depth. Differences in visual short-term memory performance appear to arise from factors that influence encoding; stereopsis is not one of those factors.

ACNKOWLEDGMENTS

I would like to extend my sincerest gratitude to my supervisor Dr. Laurie Wilcox. Thank you for all that you have done over the past two years. Your creativity, vast knowledge, encouragement, and insight have helped me tremendously throughout this process. Without your guidance, this thesis would not have been possible.

I would also like to thank the rest of my thesis committee: Dr. Mazyar Fallah, Dr. Susan Murtha, and Dr. Jacob Beck for their astute comments and questions, which incented me to assess my research from various perspectives.

Thank you to my friends, for your encouragement and for celebrating my accomplishments with me along the way, whether big or small.

Jordan, my brain-mate, your support has motivated me to do my best in all that I do. Thank you for calming my nerves when needed, and celebrating with me after each achievement. We have gotten through this together and have had endless laughs and amazing times throughout. Thank you for making this experience so enjoyable.

Finally, I would like to thank my family. To my Mom Diane and Abba Avi, thank you for encouraging me to always pursue my passions. You have given me the confidence to take on anything, knowing I have your love and support along the way. Thank you for helping me day-to-day, and for making this process so manageable. I attribute much of my success to you! To my brother, sister-in-law, and grandparents, thank you for always supporting me in both my research and non-academic endeavours. My success would not be possible without you!

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I:

GENERAL INTRODUCTION

Human memory is not simply a passive warehouse for knowledge storage but is instead in a continual state of dynamic modification (Penn, Holyoak & Povinelli, 2008). In particular, it is believed that working memory continually maintains, manipulates and updates information (Baddeley & Hitch, 1974). Working memory can be considered in terms of the input modality and its memory content (e.g. vision, audition). Research in this field has suggested that there are two distinguishable types of memory: verbal and visual memory (Luck & Vogel, 1997; Paulesu, Frith & Frackowiak, 1993). The main differences are the information they store, the encoding strategies used, and their storage duration. That is, as their names suggest, verbal memory retains verbal information and visual memory retains visual stimuli.

Verbal memory refers to the memory for words and other properties of language (e.g. memorization of letters)(Shallice & Warrington, 1970). It has been argued that verbal memory and visual memory employ different encoding strategies (Keogh & Pearson, 2011). Verbal memory can be activated by either verbal or visual stimuli. For example, to read this sentence one typically uses both visual and verbal information; in the act of reading, the visual system is being used to retain verbal content. In traditional verbal short-term memory studies, unrelated words or letters are encoded and then are immediately recalled serially. Under these conditions, memory performance appears to be highly dependent on the phonological properties of the to-be-remembered stimuli, and therefore relies on a phonological representation of stimuli (Baddeley & Hitch, 1974). However, nonphonological effects on verbal memory tasks have also been found. For example, the *lexicality effect*, where immediate recall is greater for words compared to non-words (Hulme, Maughan, & Brown, 1991), and the *frequency effect,* where high-frequency words are better remembered than low-frequency words (Hulme et al., 1997), both demonstrate the contribution of a nonphonological mechanism. When working with patients with semantic dementia, Jefferies, Frankish and Ralph (2006) demonstrated that lexicality, imaginability, word frequency, and the ratio of words to non-words, influence the strength of the phonological trace. Those patients were more likely to accurately recall word phonemes in the correct configuration, compared to nonword elements, which were either remembered in the wrong order or simply forgotten (Jefferies, Frankish & Ralph, 2006). When healthy individuals are instructed to remember a particular object, a variety of encoding strategies can be used (Hales & Brewer, 2012). For example, the word "dog" can be retained by encoding an image of a dog, by acoustically rehearsing the word dog, by encoding it semantically, etc. Further, the strategies used to encode visual vs. verbal information of the same items, result in different brain activation (Hales & Brewer, 2012). A meta-analysis of brain activation during verbal or word-based encoding showed increased activation in the left inferior frontal, superior temporal, lingual, parietal, and medial frontal cortex (Kim, 2011). These regions are distinct from the regions activated during visual encoding. Another distinction arises when comparing retrieval time; retrieval speed is more rapid for verbal, than visual, memory (Standing, 1973).

Visual memory experiments conducted over a century ago established a connection between visual representations and encoding of external visual input (Perky, 1910). Perky (1910) showed that subliminally presented visual stimuli intruded into an individual's mental images. In more recent experiments, creating a mental image of a visual stimulus, in particular an oriented line, activates early visual cortex in the same way short-term memory maintenance and visual perception of the same stimulus does (Kosslyn et al., 1999; Serences, Ester, Vogel & Awh, 2009). As well, when assessing higher level visual processing, researchers have found that neurons in face-selective regions exhibit sustained amplitude increases when an observer is retaining a face in working memory, similar to when they are looking at a face (Druzgal & D'Esposito, 2001). During visual or picture-based encoding, memory effects have been found in the superior parietal, lateral occipital, right prefrontal, and fusiform cortex (Kim, 2011). It is unclear whether the variability in brain responses between visual and verbal stimuli is due to bottom-up processing of stimuli or top-down strategy (Hales & Brewer, 2012). Unlike verbal memory encoding, where verbal information is cognitively rehearsed, visual mental representations are formed for storage of visual memories. Consistent with the proposal that there is competition for a limited amount of resources that are shared between visual processing and maintenance of a mental image, maintaining multiple items in visual memory has been shown to impair detection of concurrently presented visual targets (Farah, 1989; Konstantinou, Bahrami, Rees & Lavie, et al., 2012).

Memory can also be categorized according to storage duration, and can be separated into short-term and long-term memory (Luck & Hollingworth, 2008). The distinctions between these two types of memory are particularly noteworthy when assessing visual memory. As the name implies, visual long-term memory (VLTM) stores visual information for long periods of time (Standing, 1973). Conversely, visual short-term memory (VSTM) is a short-term buffer that stores visual information (Phillips, 1974).

Visual long-term memory does not have a known set capacity (Simons & Rensink, 2005). It contains much more detailed visual representations than VSTM (Simons & Rensink, 2005). VLTM representations develop slowly (ranging from hours to months) over time (Arganoff, Davis & Brink, 1965). Visual memories in long-term memory are maintained via passive structural modifications and therefore do not dissipate following active maintenance (Ruchkin, Grafman, Cameron & Berndt, 2003). Research conducted on goldfish suggests that protein synthesis is only invoked for long-term memory consolidation (Arganoff et al., 1965). Therefore, unlike VSTM, visual representations stored in VLTM can persist indefinitely (Brown, 1958) and contain much more visual information (Brady, Konkle, Alvarez & Oliva, 2008; Simons & Rensink, 2005).

Visual short-term memory is a cognitive process that is responsible for encoding, maintenance and retrieval of visual information (Xu & Chun, 2006). Working memory is a system that affords storage and manipulation of information needed for cognitive tasks; tasks which employ knowledge, thought processes and goal structures (Baddeley, 1992; Schraagen, Chipman & Shalin, 2000). Working memory is typically separated into three subcomponents: the central executive, the visuospatial sketchpad, and the phonological loop. Although the researchers who proposed these models once believed they opposed one another, they are now believed to describe different processes that function synergistically (Baddeley, 1986). While working memory functions during cognitive tasks, visual short-term memory acts as the visual storage component of the working memory model that is used during such tasks (Baddeley, 1986; Luck & Hollingworth, 2008). VSTM encodes information by entering visual input into storage for later recall.

Encoding is typically an active process, however certain attributes of a visual object can be stored unintentionally. Chen and Wyble (2015) showed that individuals automatically encode the 2-D location of an object, but not other object properties (e.g. colour), when they were not specifically instructed to remember such information. While encoding typically requires attention, VSTM is resilient throughout eye blinks, eye movements, and other visual interruptions, and creates largely schematic representations (Luck, 2007). Those schematic visual representations are generated very quickly (20-50ms/item) and are maintained via active mechanisms that are implemented at the neural level, requiring neural firing during the memory's retention (Ruchkin, Grafman, Cameron & Berndt, 2003; Vogel, Woodman & Luck, 2006). Because VSTM does not structurally modify the brain, once active maintenance ceases, VSTM representations are terminated (Brown, 1958).

Visual short-term memory retains only a limited amount of visual information, approximately four items (Luck & Vogel, 1997). In their influential work on VSTM capacity, Luck & Vogel (1997) designed a study that prevented contamination from verbal memory. A verbal task, storing two digits in verbal memory, was implemented at the beginning of each trial prior to the visual stimulus. The concurrent verbal memory load did not impact performance on the visual task, but it did interfere with verbal memory when the same task preceded a verbal memory task. This finding emphasizes the importance of ensuring minimal contamination from verbal memory (i.e. remembering visual tasks verbally) in visual memory experiments. The fixed capacity of four items determined by Luck & Vogel (1997), generally holds true even with the presence of conjunctive features (e.g. colour and orientation). However, Xu and Chun (2006) suggest that VSTM capacity is determined by both a fixed number of objects as well as by the complexity of the objects. It has been proposed that this set capacity is limited due to attention (Cowan, 2001). Lower VSTM capacity estimates have been attributed to dissimilar encoding content compared to the visual target; when there is a high similarity between the stimuli being encoded and tested, VSTM capacity remains roughly fixed (Awh, Barton & Vogel, 2007; Soto, Wriglesworth, Bahrami-Balani, & Humphreys, 2010). However, Van den Berg, Shin, Chou, George & Ma (2012) have argued instead that a set capacity does not exist and opposes the idea of a fixed number of slots. Their model posits that VSTM capacity is continuous and variable across items and trials, rather than discrete and fixed. They suggest that VSTM capacity should be conceptualized in terms of quality rather than number of items, and accounts for why an individual's subjective experience of VSTM can vary (Van den Berg et al., 2012).

While VSTM has been studied extensively, it has almost exclusively been directed at understanding recall of 2-D object properties. This is surprising, given that for most individuals, the visual environment is seen in three dimensions. There are many cues to depth, including both monocular and binocular depth cues. Monocular cues are those that require only one eye for the perception of depth and include pictorial, movement, and oculomotor cues (Howard & Rogers, 2012). The effectiveness of a visual cue is often dependent on the distance of the image. While occlusion, when one object hides part of another object, is useful at any distance, atmospheric perspective, a slight blue tint and decrease in sharpness for objects far away in a natural scene, is only useful at a distance greater than 30 meters (Cutting & Vishton, 1995). Binocular cues are those that require two eyes for the perception of depth. The positional difference between the images in the left and right eyes is known as binocular disparity and the perception of depth that results from binocular disparity is called stereopsis (Howard & Rogers, 2012). Considering depth cues in isolation, stereopsis produces the most compelling sense of depth and space.

The limited research on visual short-term memory that has addressed the influence of stereoscopic 3-D has produced conflicting results. In their studies of the effect of stereoscopic 3-D on VSTM, Xu & Nakayama (2007) found that stereoscopic 3-D enhanced VSTM. They argued that presenting objects on two disparity-defined surfaces, rather than one, improves VSTM performance. However, in their 3-D condition they presented their stimuli on separate depth planes which were presented sequentially; the two disparate planes were never seen simultaneously. This sequential presentation paradigm gave observers time to converge their eyes on the stimulus plane, which effectively eliminated any relative disparity. Therefore, any differences observed between conditions could not have been due to relative disparity. Instead it is likely that their results are related to general consequences of sequential object presentation for VSTM. That is, Ihssen, Liden, & Shapiro (2010) reported that VSTM performance improves when to-be-remembered items are divided into sequential arrays compared to simultaneous presentation of the same items (Ihssen, Liden, & Shapiro, 2010). While Xu and Nakayama’s (2007) results are difficult to interpret, studies of attention, a process closely linked with memory, have found that the presence of depth facilitates attention. For example, Nakayama and Silverman (1986) found that having two simultaneously presented disparity-defined planes allowed observers to attend to the different disparities serially. In addition, they could perform a search task within a given plane in parallel

(Nakayama & Silverman, 1986). The positive impact of 3-D on attention suggests that under the right circumstances, stereoscopic 3-D could be useful for VSTM; however the experiment conducted by Xu & Nakayama (2007), which found that effect, was flawed because it lacked the presence of disparity throughout the experiment. In a more recent study on VSTM, Reeves and Lei (2014) argued that separating elements across disparity-defined surfaces does not influence VSTM. However, while Reeves and Lei did present their stimuli in 3-D, they used letter stimuli, which could have been stored in verbal, rather than visual, short-term memory. Consequently, the impact of stereoscopic 3-D on visual short-term memory remains uncertain.

The purpose of my thesis is to determine if stereoscopic 3-D impacts visual short-term memory. I assess the effects of disparity and element distribution on the ability to encode and retrieve visual information. To accomplish these objectives, experiments were carefully designed to avoid the pitfalls identified in existing experiments.

Specifically, in Chapter II I begin by determining the stimulus features necessary to evaluate the role of stereoscopic 3-D on VSTM. I then assess the role of sample-test similarity on memory performance, and determine the influence of depth on VSTM when disparity is fixed. In Chapter III I evaluate the influence of stereoscopic 3-D on VSTM when disparity is varied. I also assess whether VSTM is influenced by the way objects are segregated in depth, and if observers see the depth in the stimuli clearly across the range of disparities used here. In Chapter IV, I investigate how encoding influences VSTM performance, particularly when encoding is made more difficult, versus when it is facilitated. Lastly, in Chapter V, I discuss the findings from this thesis and how they relate to existing research.

CHAPTER II:

EVALUATING VERBAL SHORT-TERM MEMORY AND SAMPLE-TEST SIMILARITY

GENERAL METHODS

Participants

Participants were recruited through the Undergraduate Research Participant Pool (URPP) at York University, as well as by word of mouth. URPP participants received credit for their participation; non-URPP participants received $10/hr. Observers' stereoacuity was assessed using the Preschool Randot Test™, and only those that could see depth from disparity of at least 40 arcmin were included in the study. Participants had normal, or corrected to normal, visual acuity. The age of participants ranged from 18 to 30 years.

Apparatus

In all experiments described here, the stimuli were presented using a modified Wheatstone stereoscope, which permitted separate presentation of images to the two eyes. Head position was stabilized using a chinrest, and the viewing distance was fixed at 55 cm. Stimuli were displayed on two CRT monitors (1600 x 1200) with a refresh rate of 100Hz. Under those conditions, one pixel subtended 1.24 arcmin. Stimulus presentation and data acquisition were controlled via CMStorm computer and purpose built Matlab code using the PsychToolbox package (Brainard, 1997; Pelli, 1997). The same apparatus was used in all of the experiments described here.

Stimuli

Letters. For each trial, a fixed number of white capital letters were quasi-randomly (omitting vowels and repeats) chosen and presented on a black background.

The font was randomized across letters and trials. The fonts used were: 'Cambria', 'Arial', 'Times New Roman', 'Comic Sans MS', and 'Century Gothic'. The letters were presented within a central rectangle outlined in white with a visual angle of 0.20 x 0.25 arcmin = 0.05 arcmin^2.

Objects. The object stimuli were presented within the central rectangle described above. The rectangle contained 24 non-overlapping square brackets; half of these were vertical (either to the left or right, but not both;] [ ) and half horizontal (either up or down, but not both; ⌙ ¬). Four conditions were tested, two 3-D conditions (*3-D mixed and 3-D segregated)* and two 2-D conditions *(2-D sparse* and *2-D dense)*. Two disparity-defined depth planes with equal element density (12) were present in both 3-D conditions. The elements were presented at zero disparity in both 2-D conditions. *3-D mixed:* the oriented elements were non-uniformly distributed across two depth planes. *3-D segregated:* the oriented elements were uniformly distributed across two depth planes (i.e. each depth plane contains only one orientation). *2-D dense:* there were 24 elements, the same as the 3-D conditions, located on a single 2-D plane *2-D sparse:* half of the number of elements (12) were displayed on a single 2-D plane. This condition was included for comparison with the 3-D mixed condition; if an observer looked at only one depth plane in the 3-D mixed condition, they would obtain similar information as the 2-D sparse condition.

Procedure

A one alternative forced choice (1AFC) same-different task was used to assess verbal short-term memory for letters and visual short-term memory for objects. In a

1AFC same-different task, a pair of stimuli (either the same or different) are presented in a trial with a delay between stimulus presentation, and observers respond whether the pair of stimuli are the same or different (Macmillan & Creelman, 2005). The letter and object tasks were described to each participant before starting the experiment, and 10 practice trials were completed. They were told to remember the set of letters (and that their order would change), as well as the two directions the square brackets were facing (one vertically oriented and one horizontally oriented). There were four possible combinations of square bracket orientations ([¬, [⌣ ,] ¬,] ⌣). Each trial began with a fixation cross (1000 ms), which was followed by a set of letters (750 ms) presented in the centre of the screen. Then a set of objects (square brackets, 1000ms) appeared in the same location. A second set of letters appeared and observers responded whether the set of letters were the same or different from the first set. The font and order of the letters were irrelevant to the task (i.e. 'J P S Z' was considered the same as 'P Z J S'); therefore observers could only use verbal short-term memory to remember the letters. Following this response, a second set of square brackets appeared. Observers indicated if the orientations of the square brackets were the same or different from the orientations of the first set of square brackets. See *Figure 1* for a depiction of the experimental sequence. All experiments use this general method and any deviations from the protocol will be noted where appropriate.

The duration of the object presentation allowed for observers to change their vergence while viewing the stimuli, however because there were always two planes offset in depth in the 3-D conditions, relative depth was always present. The stimulus duration also prevented temporary suppression of visual attention, known as the attentional blink, immediately following the presentation of the letter stimuli (Raymond, Shapiro & Arnell, 1992).

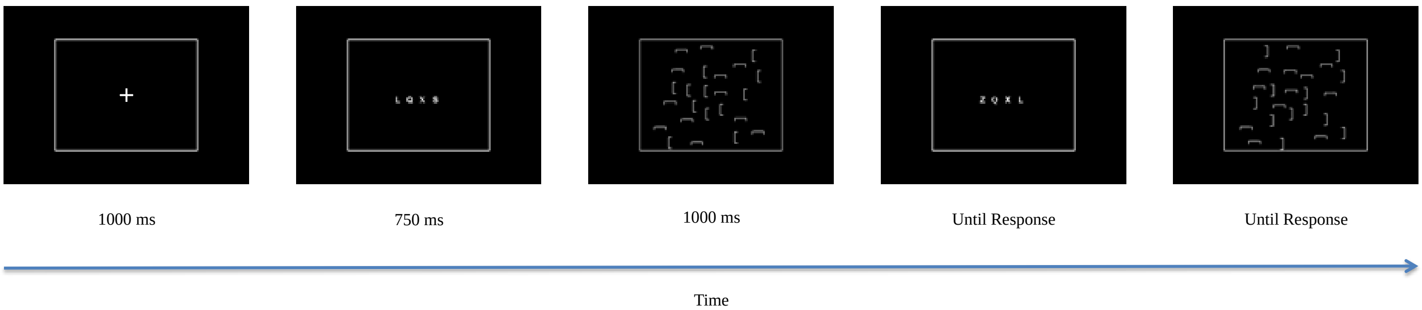

Figure 1. The general experimental timeline.

RESULTS

Sensitivity was computed by calculating d', an unbiased measure of performance, as well as proportion correct. With proportion correct alone, one cannot evaluate or discount the effects of response criterion. For instance, liberal responders may have a high 'hit rate' and therefore a high proportion correct, but will also incur a large number of false positives (Green & Swets, 1966). I used a conventional approach to compute d'; responses were coded as hits (participant responded 'same' and the test was the same as the sample), and false alarms (participant responded 'same' and the test was different from the sample). I then computed d' by subtracting the *z* score (the inverse of the normal distribution function) of the false alarms (*FA*) from the *z* score of the hits (*H*) (see *Equation 1*).

$$d' = z(H) - z(FA) \tag{1}$$

The larger the d’ value, the better the performance, with a maximum score of 4. Throughout this thesis I report d’ and for ease of explication, the associated proportion correct values are reported in Appendix A. In all experiments, data was analyzed using the statistical analysis package ‘R’.

PILOT EXPERIMENT

The main purpose of the following experiments were to determine whether VSTM is influenced by stereoscopic 3-D. Traditional VSTM experiments use 2-D stimuli; these experiments compare simple visual stimuli under 2-D and 3-D conditions. Because the focus of the study was on *visual* short-term memory, it was critical that verbal short-term memory could not be used to perform the visual task. To this end, prior to the main experiments, I conducted a pilot study to evaluate the number of letters needed to consume verbal short-term memory.

METHODS

Participants

Participants were recruited as outlined in the General Methods section. Three digit spans were tested separately: three-letter (N = 5), four-letter (N = 7), and five-letter (N = 6). Five of the participants took part in all three versions of the experiment; the others only took part in a single letter manipulation.

Stimuli

Letters. The letter stimuli are described in the General Methods section, however in this study, the number of letters varied for each version of the experiment. A set of three, four, or five letters was used for sample and test stimuli for the *three letter*, *four letter,* and *five letter* conditions respectively. The letters in the 3-D conditions were randomly presented on a depth plane that was either in front of, or behind, the fixation plane (5.2 arcmin). The letters were presented at this offset from the fixation plane for consistency with the object task. In the 3-D object conditions, the depth planes were offset by approximately 5.2 arcmin in front of, and behind, the fixation plane, for a total of 10.3 arcmin of disparity.

*Objects***.** The sample display with a full field of square brackets (at two orientations) is described in the General Methods. In these experiments, the test display contained only two white square brackets displayed on a black background in the centre of the screen, within the same rectangle as the sample stimuli. In the ‘same’ conditions, the two orientations in the test phase were the same as the two orientations in the sample phase. In the ‘different’ conditions, one of the two orientations in the test phase was different from the two orientations in the sample phase. The brackets in the 3-D conditions were presented with a constant disparity between depth planes (10.3 arcmin).

Procedure

The procedure is outlined in the General Methods section.

RESULTS AND DISCUSSION

The number of letters needed to consume verbal short-term memory was assessed and proportion correct and d’ were computed. As shown in *Figure* 2, observers’ sensitivity scores on the letter task for *three letters, four letters,* and *five letters* were compared. A repeated-measures ANOVA, comparing performance on the letters tasks for observers that took part in all three tasks (N = 5), determined a significant effect of letter quantity ($F(2,12) = 11.83$, $p = .001$). Using pairwise t-tests and the Benjamini & Hochberg (1995) method for controlling false discovery rates, the analysis showed that sensitivity decreased sequentially as the number of letters increased. Observers sensitivity was significantly higher in the *three letter* compared to *four letter* condition ($p = .004$), as well as in the *four letter* compared to *five letter* condition ($p < .001$).

The results of the pilot experiment demonstrate that a set of four letters is optimal for consuming verbal short-term memory. When there are fewer than four letters, the verbal task difficulty is too easy. Individuals are easily able to retain three letters and can encode the subsequently presented visual information using verbal strategies. When there are greater than four letters, the verbal task becomes too difficult and individuals no longer attempt to remember the entire set of letters. In this case they choose instead to only retain a few letters from the set. Once again, this allows observers to encode visual information using verbal strategies. When a set of four letters is used, individuals are able to remember the entire set of letters, but are unable to encode subsequent visual information with verbal strategies. Therefore, four letters were used for the following experiments, as they adequately consumed verbal short-term memory.

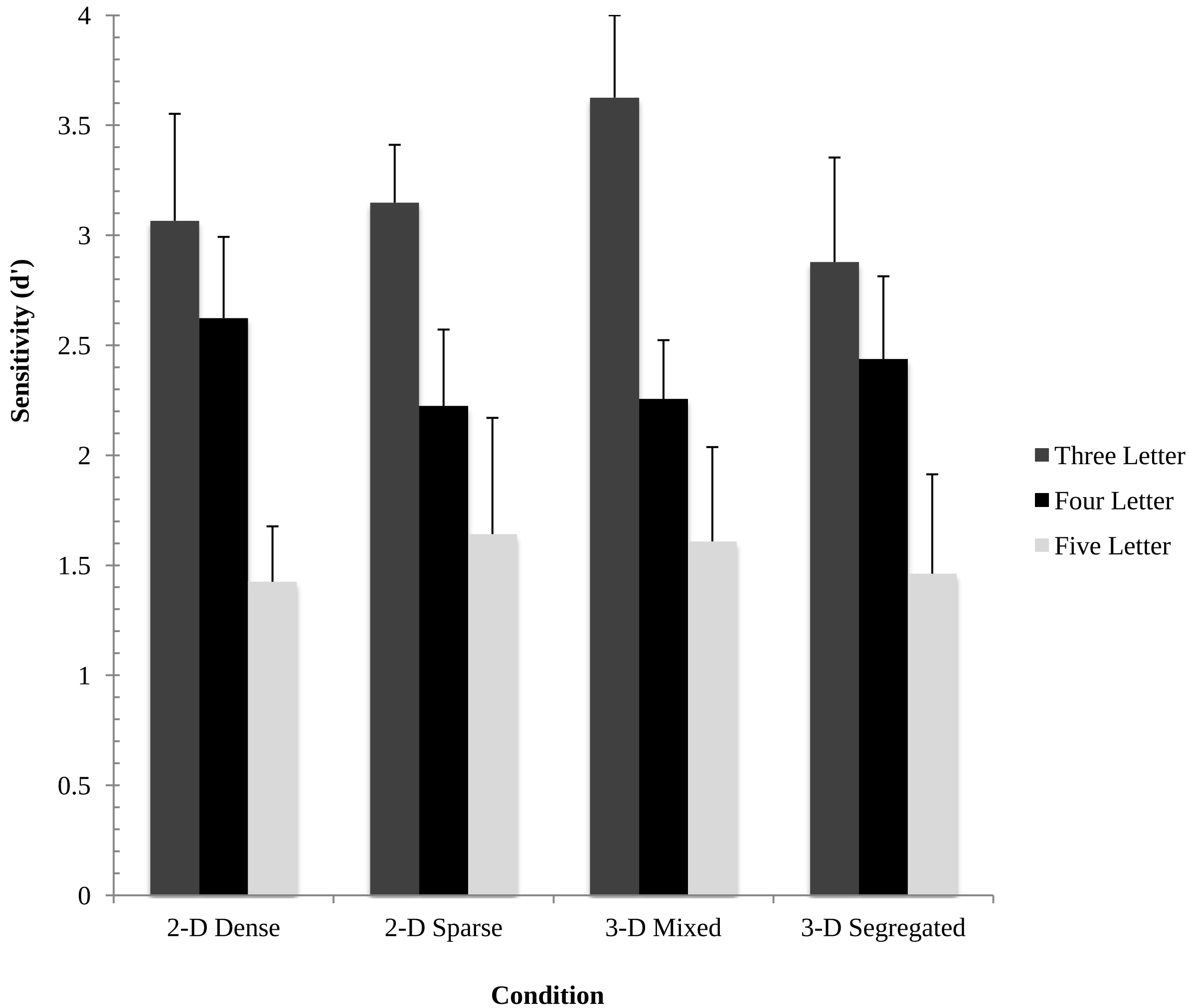

Figure 2. Letter task d' scores averaged across observers for conditions with three, four, and five letters. Error bars represent one standard error of the mean.

EXPERIMENT 1

Previous research has suggested that VSTM capacity may be influenced by the similarity between the sample information being encoded and the stimuli later being tested (Awh et al., 2007). When tasks are used that have a low similarity between the sample and test stimuli, poor performance has been attributed to mental effort being devoted to distinguishing differences between the sample and test, rather than to the memory task itself (Awh et al., 2007). While the aim of this experiment was to assess the impact of stereoscopic 3-D on VSTM, it was also important to ensure that any differences in performance were due to the variable of interest, rather than effects of sample-test similarity. Thus, I conducted both a low sample-test similarity and a high sample-test similarity experiment to assess the impact of stereopsis on VSTM. Observers' performance was analyzed independently for the two experiments to determine if VSTM sensitivity was influenced by the presence of stereoscopic 3-D at a fixed disparity. VSTM sensitivity was then compared between the two experiments to evaluate the importance of sample-test similarity.

METHODS

Participants

Participants were recruited as outlined in the General Methods section for the low sample-test similarity experiment (N=16) and high sample-test similarity experiment (N = 15). Due to time constraints, only twelve of those participants took part in both the low and high sample-test similarity experiments.

Stimuli

*Letters***.** The letter stimuli are described in the General Methods, and as outlined above, four letters were used to consume verbal short-term memory. The letters in the 2-D conditions were presented with zero disparity (0 arcmin), and the letters in the 3-D conditions were randomly presented on a depth plane that was either in front of, or behind, the fixation plane (5.2 arcmin).

*Objects***.** To evaluate the influence of the similarity between the encoded sample objects and the test objects on VSTM, two experiments were conducted with the same sample stimuli, but different test stimuli. *Low sample-test similarity*. A full field of square brackets was shown in the encoding phase of the experiment, as outlined in the general procedure. The positions of the square brackets in the encoding phase were randomized on each trial. However, for the test phase, only two square brackets were shown. In the 'same' conditions, the two square brackets in the test display had the same orientations as the square brackets shown in the sample display. In the 'different' conditions, one of the two square brackets shown in the test display had a different orientation from those shown in the sample display. *High sample-test similarity.* Both the sample and test objects consisted of a full field of square brackets. The positions of the square brackets were held constant on each trial for the sample and test, however the orientations of the square brackets were randomly distributed across those locations. In the 'same' conditions, the two orientations were the same, but at any given position, the orientation was chosen randomly (from the two options). In the 'different' conditions, one of the two orientations was different. For example, if the sample display had 12 leftward facing brackets and 12 upward facing brackets, a potential *different* condition would be when all of the upward facing square brackets faced down in the test phase (at any given position, the orientation was chosen at random; 12 left, 12 down). Consistent with the *letter task*, the square brackets in the 2-D conditions were presented with zero disparity (0 arcmin), and the square brackets in the 3-D conditions were presented with a disparity of 10.3 arcmin between planes.

Procedure

The procedure is outlined in the General Methods section. The *low sample-test similarity* and *high sample-test similarity* experiments were tested separately. The order of the two experiments was chosen quasi-randomly for each participant. Within each experiment there were five repetitions of the square bracket orientation combinations for each condition (5 repetitions x 4 square bracket orientation combinations x 4 conditions) resulting in 80 trials per experiment. Performance on each experiment was evaluated separately; as well, performance between the two experiments was compared for observers that completed both.

RESULTS AND DISCUSSION

As shown in *Figure 3*, within the low sample-test similarity and high sample-test similarity experiments, observers performed consistently regardless of whether or not there was disparity between depth planes. Proportion correct graphs for all experiments can be found in Appendix A. A repeated measures ANOVA was used to compare observer's performance in the low sample-test similarity experiment, and revealed no significant difference in sensitivity between the 2-D and 3-D conditions ($F(1,15) = 1.80$, *p* = .20). Using BH adjusted pairwise t-tests, there was no significant difference in sensitivity between the 3-D mixed and 3-D segregated conditions (*t*(15) = 0.23, *p* = 0.82). A repeated measures ANOVA was used to evaluate observer's performance in the high sample-test similarity experiment, and revealed that there was also no significant difference in sensitivity between the 2-D and 3-D conditions (F(1,14) = 0.39, *p* = .54). Using BH adjusted pairwise t-tests comparing the 3-D conditions, there was no significant difference in sensitivity between the mixed and segregated conditions (*t*(14) = -1.35, *p* = 0.20).

There was however a difference in VSTM performance based on sample-test similarity, as shown in *Figure 4*. The impact of the similarity between the sample stimuli and the test stimuli was compared using a t-test for planned comparisons. Observer's were significantly more sensitive to stimuli when there was a high similarity between the sample and test stimuli compared to when there was a low similarity between the sample and test stimuli (*t*(47) = -2.41, *p* =0.02).

The results of this experiment demonstrated that there was no difference in visual short-term memory performance when objects were viewed in 2-D or 3-D, regardless of the similarity between the sample and test displays. Nor was there a difference based on how objects were segregated (mixed or segregated) in the 3-D display. The results of Experiment 1 did however demonstrate that individuals perform better on visual memory tasks when there is a higher similarity between the information being encoded and the information being tested, consistent with findings by Awh et al. (2007). The fact that performance was worse when the test was different from the sample, suggests that individuals were focusing on irrelevant aspects of the memory task (i.e. distinguishing differences in the sample and test stimuli), which detracted from their performance. To avoid this potential confound, in subsequent experiments I use the high sample-test similarity display paradigm.

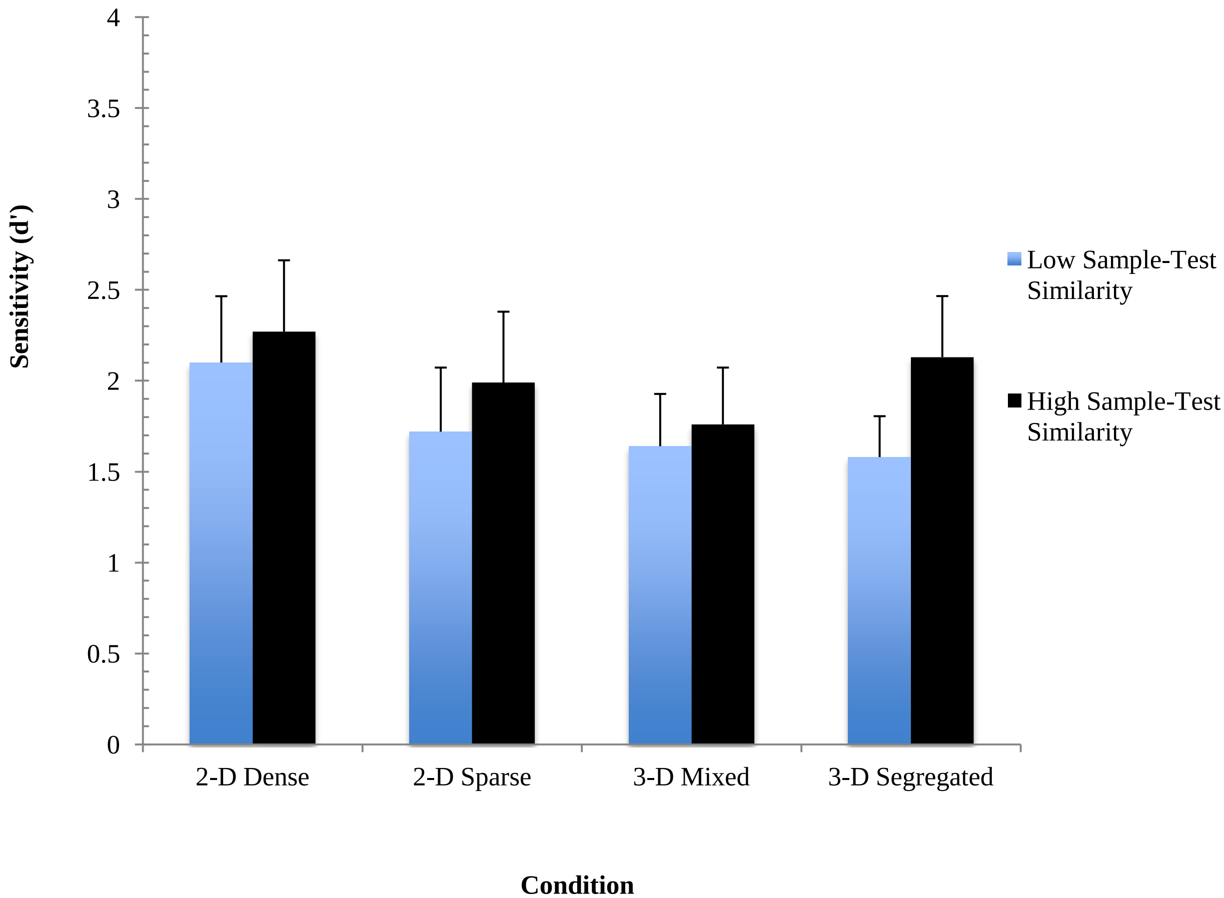

Figure 3. Object task d’ scores averaged across observers for the low sample-test similarity (N = 16) and high sample-test similarity (N = 15) experiment. Error bars represent one standard error of the mean.

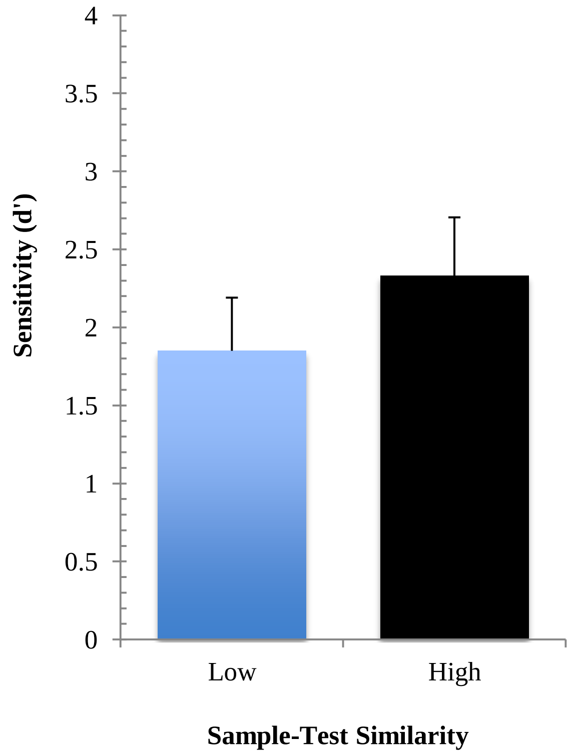

Figure 4. Object task d’ scores averaged for observers that took part in both the low and high sample-test similarity experiments (N = 12). Error bars represent one standard error of the mean.

CHAPTER III:

THE IMPACT OF SEGREGATING OBJECTS IN STEREOSCOPIC 3-D ON VISUAL SHORT-TERM MEMORY

EXPERIMENT 2

In Experiment 1, I showed there was no difference in visual short-term memory performance regardless of whether the visual stimuli were presented in 2-D or 3-D. While it appears there is no impact of distributing elements in 3-D on VSTM, in that study I used a fixed disparity of 10.3 arcmin, as participants reported in preliminary studies that this amount of disparity was comfortable and provided a stable percept of depth. However, it is possible that the lack of effect reported in Experiment 1 was tied to the specific test disparity used. In Experiment 2, a range of disparities was used to re-evaluate the impact of disparity on VSTM sensitivity. This experiment was split into two parts, one experiment where the 3-D conditions (mixed and segregated) were interleaved, and another where the 3-D conditions were tested in separate blocks.

EXPERIMENT 2A

The results of Experiment 1 showed that there was no difference in VSTM when objects were displayed in 2-D or 3-D. In the 3-D condition of that experiment, a fixed amount of disparity (10.3 arcmin) was used, and the mixed and segregated conditions were interleaved. To determine if the amount of disparity plays a role in VSTM, a range of disparities (6.4 arcmin, 12.9 arcmin, 19.3 arcmin, and 25.8 arcmin) was selected that bracketed the amount of disparity used in Experiment 1. Keeping consistent with the first experiment, to allow for similar encoding strategies, the mixed and segregated conditions were interleaved.

METHODS

Participants

Participants were recruited as outlined in the General Methods section (N = 21). None of these participants took part in the other experiments.

Stimuli

Letters. The 2-D conditions are described in the General Methods section. In the 3-D conditions, the set of letters were randomly presented either in front of, or behind, the fixation plane by a varying amount of disparity on each trial (3.2 arcmin, 6.5 arcmin, 9.7 arcmin, and 12.9 arcmin).

Objects. The 2-D conditions are described in the General Methods section. The amount of disparity between depth planes varied on each trial for the 3-D conditions (6.4 arcmin, 12.9 arcmin, 19.3 arcmin, and 25.8 arcmin).

Procedure

The procedure outlined in the General Methods section was used here.

2-D conditions. The 2-D conditions (2-D sparse and 2-D dense) were tested in one block with five replications of bracket orientation combinations (5 replications x 4 square bracket orientation combinations x 2 conditions) resulting in 40 trials.

3-D conditions. To avoid fatigue, data was collected in two blocks of trials containing both the 3-D mixed and 3-D segregated conditions. One block contained two replications of the four bracket orientation combinations per disparity for both 3-D conditions (2 replications x 4 combinations of square brackets x 4 disparities x 2 conditions) resulting in 64 trials; the other block contained three replications of the four bracket orientation combinations per disparity for both 3-D conditions (3 replications x 4 combinations of square brackets x 4 disparities x 2 conditions) resulting in 96 trials. During a block, participants were given two opportunities to take a short break from testing after 21 trials and 42 trials for the first block; and after 32 trials and 64 trials for the second block.

RESULTS AND DISCUSSION

Consistent with the results of Experiment 1, visual short-term memory performance was unaffected by the presence of stereoscopic depth (see *Figure 5*). A one-way ANOVA was used to compare VSTM performance for 2-D and 3-D object stimuli when the 3-D conditions were interleaved. No significant difference was found between the 2-D and 3-D conditions ($F(1,20) = 0.01$, $p = .91$). A two-way ANOVA was used to analyze the 3-D conditions only. There was no significant difference between the 3-D mixed and 3-D segregated conditions ($F(1,20) = 0.54$, $p = .47$), nor was there an effect of disparity ($F(3,60) = 1.53$, $p = 0.22$). As well, there was no significant interaction between the nature of the 3-D condition and disparity ($F(3,60) = 2.10$, $p = 0.11$). Consistent with the previous experiment, stereoscopic 3-D does not influence VSTM sensitivity.

Individuals appear to be poorer at remembering visual information overall when the amount of depth, and the way objects are segregated, are continually being varied, compared to when a fixed amount of depth is used (see Experiment 1). However, VSTM performance is consistent regardless of whether the information is presented with or without depth. This is consistent with the results of Experiment 1, in that visual objects are remembered equally irrespective of whether they are displayed in 2-D or 3-D.

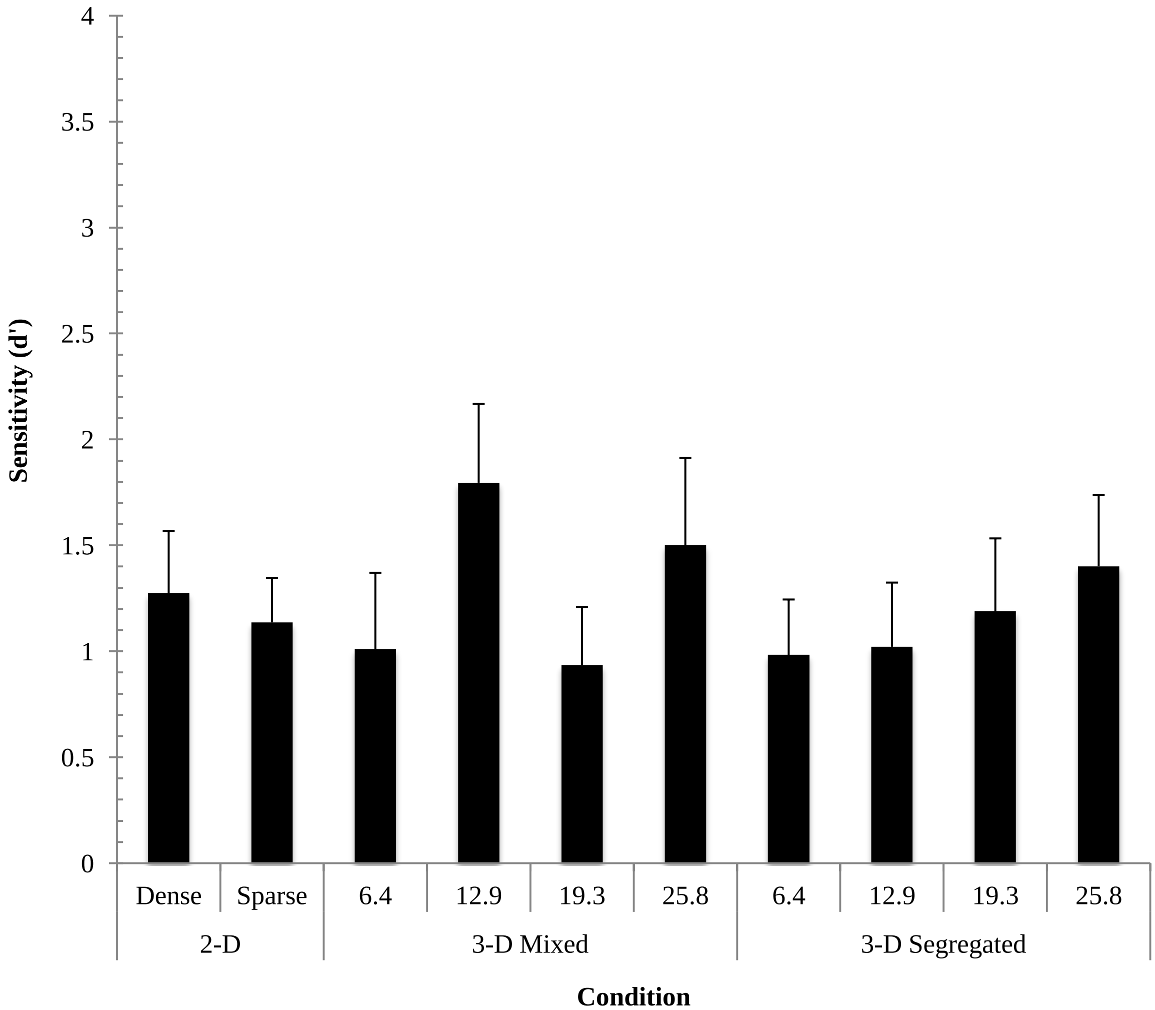

Figure 5. Object task d’ scores averaged across observers when disparity (arcmin) varied and 3-D conditions were interleaved (N = 21). Error bars represent one standard error of the mean.

EXPERIMENT 2B

In Experiment 2a, there were no differences in VSTM performance in the 2-D and 3-D conditions when disparity was varied in the 3-D conditions. However, in Experiment 2a, the mode of segregation (mixed and segregated) was interleaved. Interleaving the 3-D conditions could have prevented observers from applying consistent grouping strategies. In this experiment, disparity was still varied across trials, however the 3-D conditions were separated into blocks (3-D mixed block and 3-D segregated block). Should an observer choose to use a grouping related encoding strategy (e.g. in the 3-D segregated condition, grouping all square brackets on a depth plane), the same strategy could be used across a block of trials. The same grouping strategy would be more difficult to use when the 3-D conditions were interleaved, like in Experiment 2a, because on any given trial the square brackets may have been mixed or segregated.

METHODS

Participants

Participants were recruited as outlined in the General Methods section (N = 12). None of these participants took part in the other experiments.

Stimuli

Letters. As in Experiment 2a, the 2-D conditions are described in the General Methods section. In the 3-D conditions, the set of letters were randomly presented either in front of, or behind, the fixation plane by a varying amount of disparity on each trial (3.2 arcmin, 6.5 arcmin, 9.7 arcmin, and 12.9 arcmin).

Objects. The 2-D conditions are described in the General Methods section. The amount of disparity between depth planes varied on each trial for the 3-D conditions (6.4 arcmin, 12.9 arcmin, 19.3 arcmin, and 25.8 arcmin).

Procedure

The procedure is outlined in the General Methods section.

2-D conditions. The 2-D conditions (2-D sparse and 2-D dense) were tested in one block with five replications of bracket orientation combinations (5 replications x 4 square bracket orientation combinations x 2 conditions) resulting in 40 trials.

3-D conditions. The 3-D conditions (3-D mixed and 3-D segregated) were not interleaved and were divided into two blocks. The 3-D mixed block contained five replications of the four bracket orientation combinations (i.e. 5 replications x 4 combinations of square brackets x 4 disparities) resulting in 80 trials. The same was done for the 3-D segregated block. During a block, participants were given two opportunities to take a short break from testing after 26 trials and 53 trials.

RESULTS AND DISCUSSION

As shown in *Figure 6*, once again, visual short-term memory performance was unaffected by the presence of stereoscopic 3-D. A one-way ANOVA was used to compare visual short-term memory sensitivity for 2-D and 3-D object stimuli when 3-D conditions were tested in separate blocks (See *Figure 6*). No significant difference was found between 2-D and 3-D conditions ($F(1,12) = 0.53$, $p = .48$). A two-way ANOVA was used to analyze the 3-D conditions only. There was no significant difference between the 3-D mixed and 3-D segregated conditions ($F(1,12) = 4.52$, $p = .055$), nor was there an effect of disparity ($F(3,36) = 0.43$, $p = 0.73$). As well, there was no significant interaction between 3-D conditions and disparities ($F(3,36) = 2.14$, $p = 0.11$).

When the amount of disparity is varied across trials and 3-D conditions are blocked based on mode of segregation (mixed and segregated), individual's short-term memory performance appears to be better overall (see sensitivity scores in *Figure 6* compared to *Figure 5* overall). This suggests that individuals may use different strategies to retain visual information based on the mode of segregation, and that it is easier to use each of these strategies when the tasks are performed separately, rather than interleaved (see Experiment 2a). Although the task appears to be easier overall, there is no difference in VSTM performance regardless of whether information is presented in 2-D or 3-D. This once again shows that distributing elements across stereoscopic depth planes does not impact VSTM sensitivity.

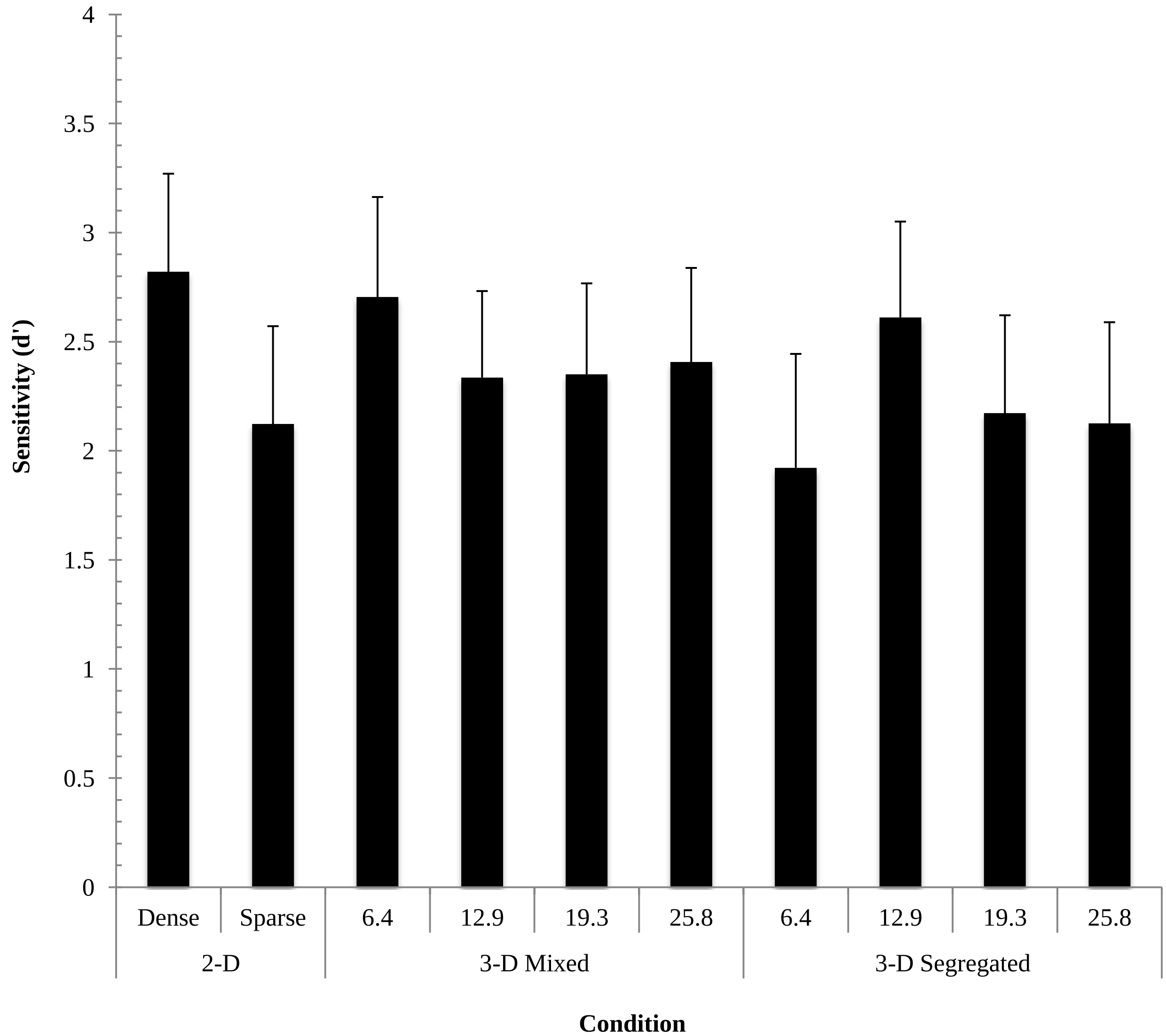

Figure 6. Object task d’ scores averaged across observers when disparity (arcmin) varied and 3-D conditions were blocked (N = 12). Error bars represent one standard error of the mean.

EXPERIMENT 3

Although there was no effect of disparity on VSTM in Experiment 2, that may have been because the amount of disparity varied from trial-to-trial. Should observers have a particular encoding strategy for a specific amount of depth between depth planes, varying disparity may have interfered with that strategy. Therefore, as in Experiment 1, in this experiment I fixed the disparity in a given block, but unlike Experiment 1, I used a smaller (6.4 arcmin) and larger (19.3 arcmin) disparity. This made it feasible for observers to use a consistent depth-related encoding strategy, should they want to.

METHODS

Participants

Participants were recruited as outlined in the General Methods section (N = 14). None of these participants took part in the other experiments.

Stimuli

Letters. Given the consistency of the 2-D results in Experiments 1, 2a and 2b, only 3-D conditions were tested in this experiment. The letters were randomly presented either in front of, or behind, the fixation plane (3.2 arcmin and 9.7 arcmin).

Objects. Consistent with the letter stimuli, there were no 2-D conditions. The 3-D conditions (3-D mixed and 3-D segregated) were interleaved, but were separated into two blocks for two disparities (6.4 arcmin and 19.3 arcmin).

Procedure

The procedure is outlined in the General Methods section. The test order of the blocks (block A: both 3-D conditions with 6.4 arcmin between depth planes, block B: both 3-D conditions with 19.3 arcmin between depth planes), were quasi-randomly selected for each participant. Each block consisted of 5 repetitions of 4 bracket orientation combinations by two 3-D conditions (i.e. 5 repetitions x 4 bracket combinations x 2 conditions) resulting in 40 trials per disparity.

RESULTS AND DISCUSSION

Observers performed consistently regardless of the amount of depth present, or the mode of segregation (see *Figure 7*). A repeated measures ANOVA was used to compare VSTM sensitivity for the two test disparities (6.4 arcmin and 19.3 arcmin). There was no significant difference between the 3-D mixed and 3-D segregated conditions ($F(1,13) = 0.04$, $p = 0.83$) nor was there a significant effect of disparity ($F(1,13) = 0.02$, $p = .88$). There was no interaction between 3-D conditions and disparities ($F(1,13) = 0.04$, $p = .85$). As observed in the previous experiments, the amount of disparity between depth planes did not influence VSTM.

Whereas the disparity varied across trials in Experiment 2a and 2b, in this experiment, the disparity was consistent from trial to trial within a block, similar to Experiment 1. Although in Experiment 2a, varying the mode of 3-D segregation (mixed vs. segregated) across trials negatively impacted performance overall, this was not the case in this experiment. The lack of impairment on performance can be explained by the consistency of the amount of depth across trials. Varying the mode of segregation was more detrimental when the amount of depth was also varied. As long as either the mode of segregation, or the amount of depth, was the same across trials, performance was better overall. Poor performance when both disparity and mode of segregation vary can likely be explained by encoding strategies. When disparity is constant throughout a set of trials, observers can use depth related encoding strategies. When mode of segregation is constant throughout a set of trials, observers can use grouping related strategies. However, when disparity and mode of segregation are both varied across trials, it is more difficult to use either strategy.

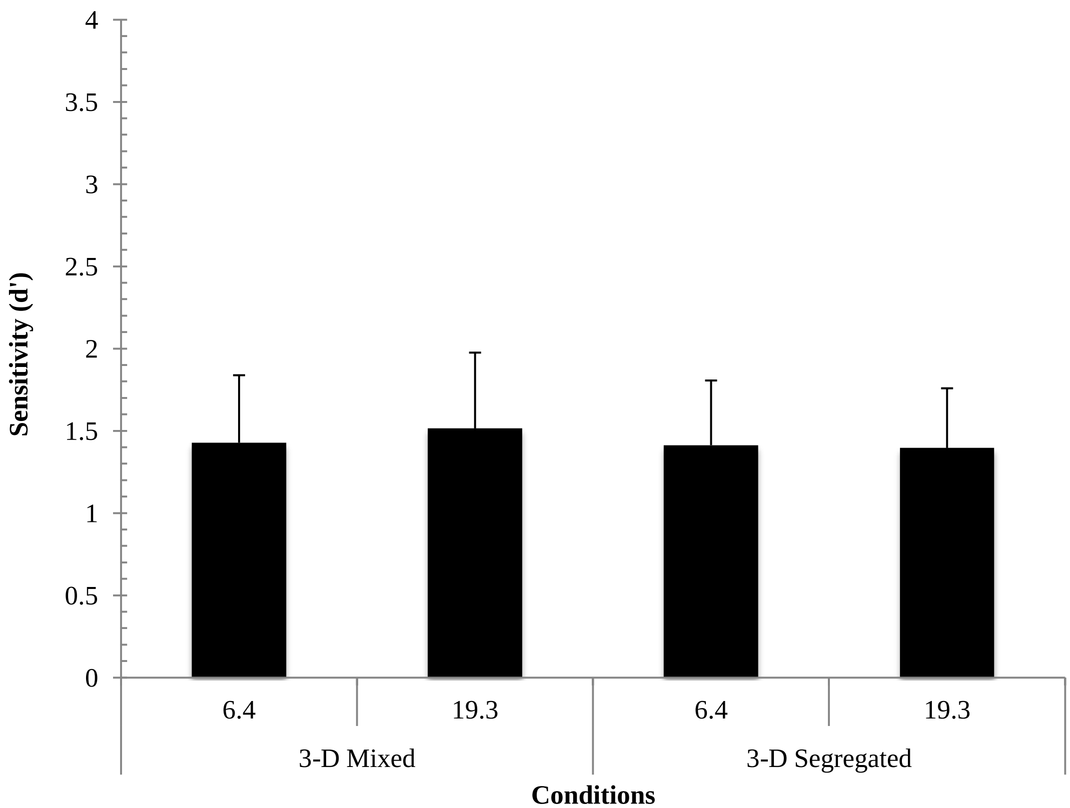

Figure 7. Object task d' scores averaged across observers when 3-D conditions were interleaved but were blocked based on disparity (arcmin) (N = 14). Error bars represent one standard error of the mean.

EXPERIMENT 4

In Experiments 1-3, I consistently reported that there is no reliable impact of disparity on VSTM. A disparity of 10.3 arcmin was chosen for the Pilot Experiment and Experiment 1 because it was comfortable for observers (assessed informally in preliminary studies). A range of disparities that bracketed the initial disparity (6.4 arcmin, 12.9 arcmin, 19.3 arcmin, and 25.8 arcmin) was then used to determine if the size of disparity influenced VSTM in further experiments. It is possible, albeit unlikely, that the larger disparities in this range were difficult to fuse or even appeared diplopic to some observers. If so they may have appeared indistinct or unclear to the naïve observers who participated in these experiments. If the stimuli were diplopic or unclear, poor VSTM performance could be attributed to an inability to properly see the stimuli, rather than an inability to encode and remember them.

METHODS

Participants

Participants were recruited as outlined in the General Methods section (N = 14).

Stimuli

Objects. A full field of square brackets was shown to observers on a given trial, as described in the general methods section. The amount of disparity between depth planes varied on each trial (6.4 arcmin, 12.9 arcmin, 19.3 arcmin, and 25.8 arcmin).

Procedure

At the start of each trial, a fixation-cross appeared (1000 ms) followed by a full field of square brackets (1000 ms). A horizontal white line then appeared with the words '*clear'* located on the top left of the line, and 'un*clear'* on the top right of the line. A shorter white vertical bar bisected the horizontal line, and observers used a game controller to shift the vertical bar towards either *clear* or *unclear* based upon how they perceived the stimuli. They then pushed a button to proceed to the next trial. Observers were told to report '*clear*' when the stimuli looked crisp and clear, and shift the bar towards '*unclear*' when the stimuli appeared fuzzy or less crisp. This experiment was conducted in 3 blocks: 3-D mixed (4 square bracket orientation combinations x 4 repetitions x 4 disparities = 64 trials), 3-D segregated (4 square bracket orientation combinations x 4 repetitions x 4 disparities = 64 trials), 2-D sparse and 2-D dense (4 square bracket orientation combinations x 4 repetitions x 2 conditions = 32 trials). Responses were coded by dividing the horizontal line such that the far left end, representing *clear*, was given a score of 0 and the far right end, representing *unclear*, was given a score of 1.

RESULTS AND DISCUSSION

As shown in *Figure 8*, there was a relationship between clarity ratings and stimulus disparity, but the difference was constrained to the small disparity range. This observation was supported by a regression analysis that suggests disparity predicts clarity at the low disparity range ($\beta = .02$, $t(68) = 6.789$, $p < .001$). Disparity also explained a significant proportion of variance in clarity ($R^2 = .40$, $F(1,68) = 46.1$, $p < .001$). It is clear from BH adjusted pairwise post-hoc t-tests of means that the effect of disparity was limited to the difference between the 2-D (zero disparity) and smallest 3-D disparity offset ($p < .001$). There was no relationship between perceived clarity and disparity for any of the remaining 3-D conditions (*p's* > .05).

The results of this experiment demonstrated that observers could properly see the stimuli when disparity was present, regardless of the amount of depth. It was likely there was a bias to remain near the middle of the horizontal line, which was evident as observers never selected either clear or unclear for the stimuli when it was in 2-D or 3-D. There was no difference in clarity between the disparities, suggesting that the large disparities were fused to form a single object. Therefore, any effects of disparity in the previous experiments were not due to large disparities being seen as double.

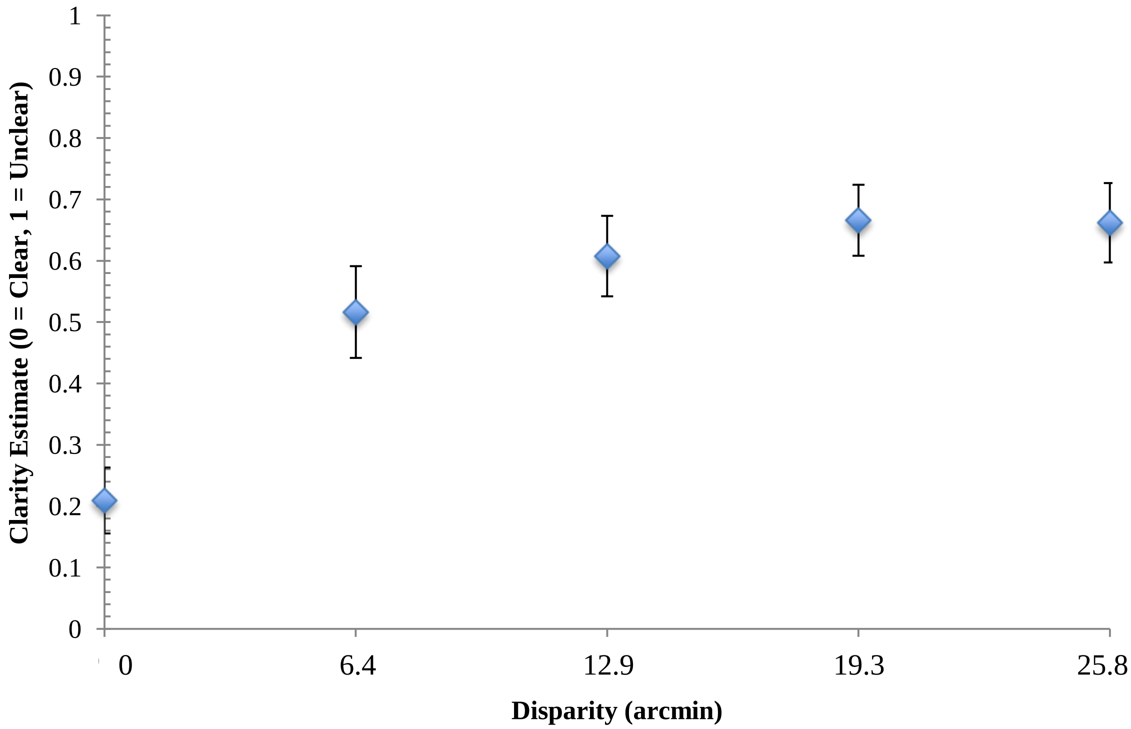

Figure 8. Observers' perception of clarity averaged across a range of disparities (N = 14). Error bars represent one standard error of the mean.

CHAPTER IV:

THE ROLE OF ENCODING

EXPERIMENT 5

In Experiments 1-3, I reported that disparity does not reliably impact visual short-term memory. It is therefore unlikely that depth-related grouping cues facilitate encoding. It is possible however, that observers use grouping strategies to help encode visual stimuli, and that these influences interfere with any advantage provided by binocular disparity. To determine whether observers were using grouping strategies to encode the orientations of the square brackets, a task was designed that would prevent such grouping strategies. Distracting elements that would impede the grouping of square brackets, were interspersed within the stimuli, while varying the amount of disparity. If observers perform worse under these conditions, it would suggest that they could no longer use grouping strategies. If observers performed better, it would suggest that observers chose to use depth-related encoding strategies to compensate for an inability to use grouping strategies.

METHODS

Participants

Participants were recruited as outlined in the General Methods section (N = 18).

Stimuli

Letters. The 2-D conditions are described in the General Methods section. In the 3-D conditions, the set of letters were randomly presented either in front, of or behind, the fixation plane by a varying amount of disparity on each trial (3.2 arcmin, 6.5 arcmin, 9.7 arcmin, and 12.9 arcmin).

Objects. A full field of square brackets was used for the sample and test stimuli, as outlined in the General Methods section. In addition, 10 diagonal lines (half oriented 45° towards the left and half oriented 45° towards the right) were placed in random positions within the rectangle, but did not overlap with the square brackets. In the 2-D conditions, the square brackets and diagonal lines were presented with zero disparity (0 arcmin). In the 3-D conditions, the amount of disparity between depth planes varied on each trial (6.4 arcmin, 12.9 arcmin, 19.3 arcmin, and 25.8 arcmin), and the distracting elements were evenly (5 on each plane) but randomly (orientation was chosen at random) distributed between depth planes.

Procedure

The procedure is outlined in the General Methods section.

2-D conditions. The 2-D conditions (2-D sparse and 2-D dense) were tested in one block with five replications of bracket orientation combinations (5 replications x 4 square bracket orientation combinations x 2 conditions), resulting in 40 trials.

3-D conditions. The 3-D conditions (3-D mixed and 3-D segregated) were interleaved and separated into two blocks. One block contained two replications of the four bracket orientation combinations per disparity for both 3-D conditions (2 replications x 4 combinations of square brackets x 4 disparities x 2 conditions) resulting in 64 trials; the other block contained three replications of the four bracket orientation combinations per disparity for both 3-D conditions (3 replications x 4 combinations of square brackets x 4 disparities x 2 conditions) resulting in 96 trials. Similar to Experiment 2, participants were given two opportunities to take a short break from testing after 21 trials and 42 trials for the shorter block; and after 32 trials and 64 trials for the longer block, to avoid attention lapses due to fatigue.

RESULTS AND DISCUSSION

As shown in *Figure 9*, the presence of distracting elements interspersed within the object display resulted in lower VSTM sensitivity. A repeated measures ANOVA was used to compare VSTM sensitivity for 2-D and 3-D object stimuli when distracting elements were present and the amount of disparity varied across trials. No significant difference was found between the 2-D and 3-D conditions ($F(1,17) = 0.0003$, $p = .98$). A two-way ANOVA was used to analyze the 3-D conditions only. There was no significant difference between the 3-D mixed and 3-D segregated conditions ($F(1,17) = 0.03$, $p = .87$), or between the different disparities ($F(3,51) = 1.21$, $p = 0.32$). However, there was a significant interaction between the 3-D conditions and the amount of disparity ($F(3,51) = 4.77$, $p = .005$). Using BH adjusted pairwise t-tests, no significant differences were found between the 3-D conditions at any disparity.

Sensitivity scores for this experiment appeared lower than other experiments, which can be explained by the increased encoding difficulty due to the presence of the distracting elements. As in the previous experiments, there was no difference between VSTM sensitivity in the 2-D compared with the 3-D conditions. Although an interaction was found between how elements were distributed across depth planes (mixed vs. segregated) and amount of disparity, no differences were found in the paired comparisons analysis. This is likely due to the high variance across subjects caused by the increased task difficulty. Overall, it appears that preventing grouping strategies resulted in lower

VSTM performance. However, even when grouping strategies were disrupted, there were still no differences in VSTM performance between 2-D and 3-D conditions. This shows that depth related encoding strategies were not used to compensate for the lack of grouping strategies.

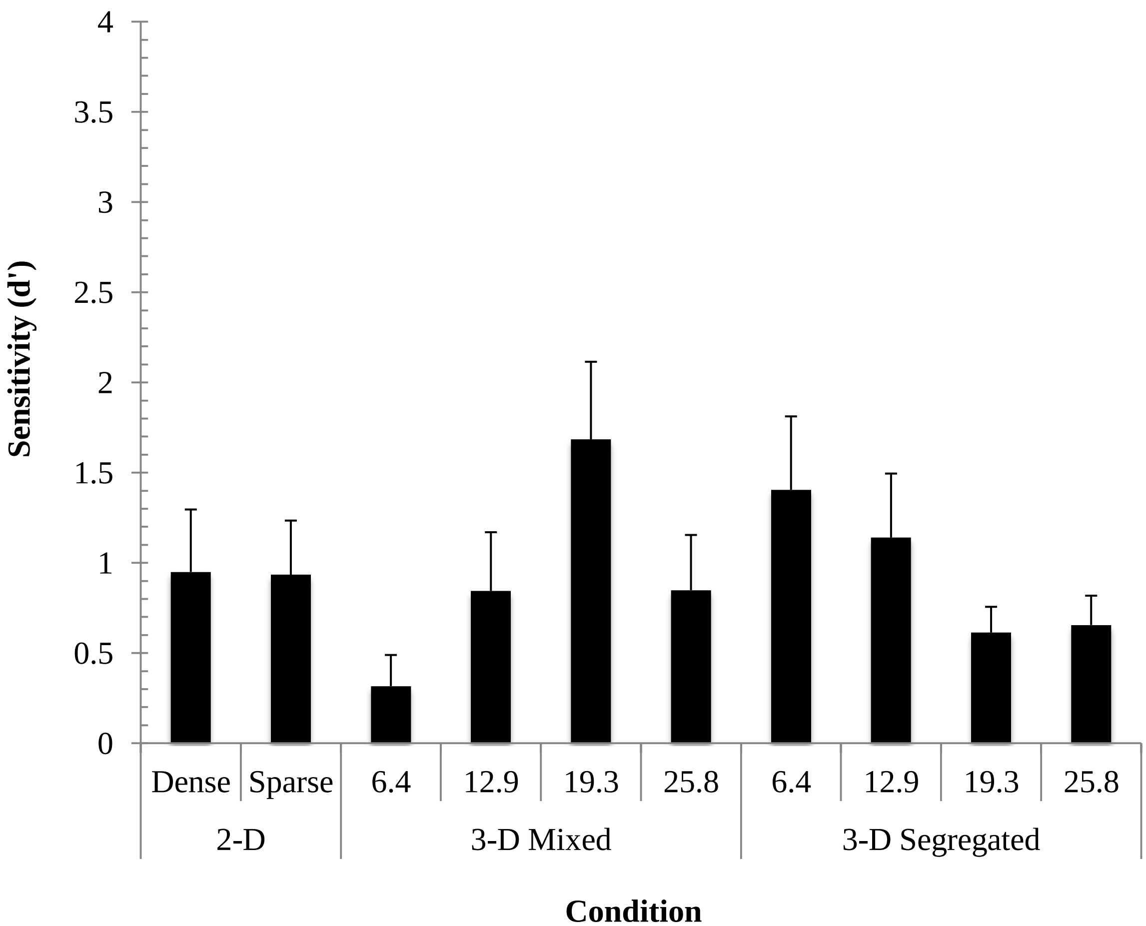

Figure 9. Object task d' scores averaged across observers when disparity (arcmin) varied and distracting elements were present (N = 18). Error bars represent one standard error of the mean.

EXPERIMENT 6

Throughout the previous experiments, I have consistently demonstrated that disparity does not improve VSTM. Overall, observers performed most poorly when encoding the object stimuli was most difficult, which emphasizes the importance of the encoding phase (see Experiment 5) irrespective of the availability of stereoscopic depth. In Experiment 5, the presence of distracting elements effectively disrupted grouping related encoding strategies; however, while the task became more difficult it did not lead to the use of depth-related encoding strategies. Albeit unlikely, it is possible that the encoding process was simply too difficult in all of the preceding experiments for 3-D segregation information to be useful. In this experiment I explicitly make it easier for observers to encode the stimuli by increasing viewing time. If task difficulty limits the use of 3-D in my previous experiments, then substantially increasing the exposure duration should lead to improvements in the 3-D conditions.

METHODS

Participants

Participants were recruited as outlined in the General Methods section (N = 10).

Stimuli

***Letters*.** The 2-D conditions are described in the General Methods section. In the 3-D conditions, the set of letters were randomly presented either in front of, or behind, the fixation plane by a varying amount of disparity on each trial (3.2 arcmin and 9.7

arcmin). The sample letters were presented for the same duration as all other experiments (750 ms).

Objects. A full field of square brackets was used for the sample and test object stimuli. In the 3-D conditions, the amount of disparity between depth planes varied on each trial (6.4 arcmin and 19.3 arcmin). Unlike prior experiments, the square brackets were presented for 2000 ms during the encoding sample phase. As with the preceding experiments, observers were given an unlimited amount of time to respond during the test phase.

Procedure

The procedure is outlined in the General Methods section.

RESULTS AND DISCUSSION

Despite facilitating the encoding phase, there were no differences in VSTM performance when stimuli were presented in 2-D or 3-D (see *Figure 10*). A repeated measures ANOVA compared VSTM sensitivity for 2-D and 3-D object stimuli when the encoding time for the object stimuli was increased (See *Figure 10*). There was no significant difference between the 2-D and 3-D conditions $F(1,9) = 0.02$, $p = .90$). A repeated measures ANOVA was used to analyze the 3-D conditions only. There was no significant difference between the 3-D mixed and 3-D segregated conditions ($F(1,9) = 3.81$, $p = .08$), or between the 6.4 arcmin and 19.3 arcmin disparities ($F(1,9) = 0.97$, $p = 0.35$) disparities. There was a marginally significant interaction between 3-D conditions and disparities ($F(1,9) = 5.36$, $p = 0.05$). At 19.3 arcmin of disparity, a BH adjusted pairwise t-test, determined there was a marginally significant difference between the 3-D mixed and 3-D segregated conditions (p = .054).

The results of Experiment 5 show that the manipulation was successful in influencing encoding; VSTM performance improved overall. However, once again, there was no difference in VSTM performance when viewing information in 3-D compared to 2-D. There was a borderline significant interaction between mode of segregation and amount of disparity, such that at larger disparities observers performed better when the visual information was non-uniformly segregated across depth. This may hint at an underlying strategy that can be used when encoding is easier. That is, when the two planes are widely separated in depth, it is easier to have non-uniformly segregated stimuli because the task can be performed by attending to a single depth plane. It appears that the main factor influencing performance here (and in the preceding experiments) is the ease of encoding. Conditions that facilitate encoding (i.e. increased encoding time) rather than make encoding more difficult (i.e. distracting elements), result in improved performance. However, the presence of stereoscopic depth planes do not benefit VSTM.

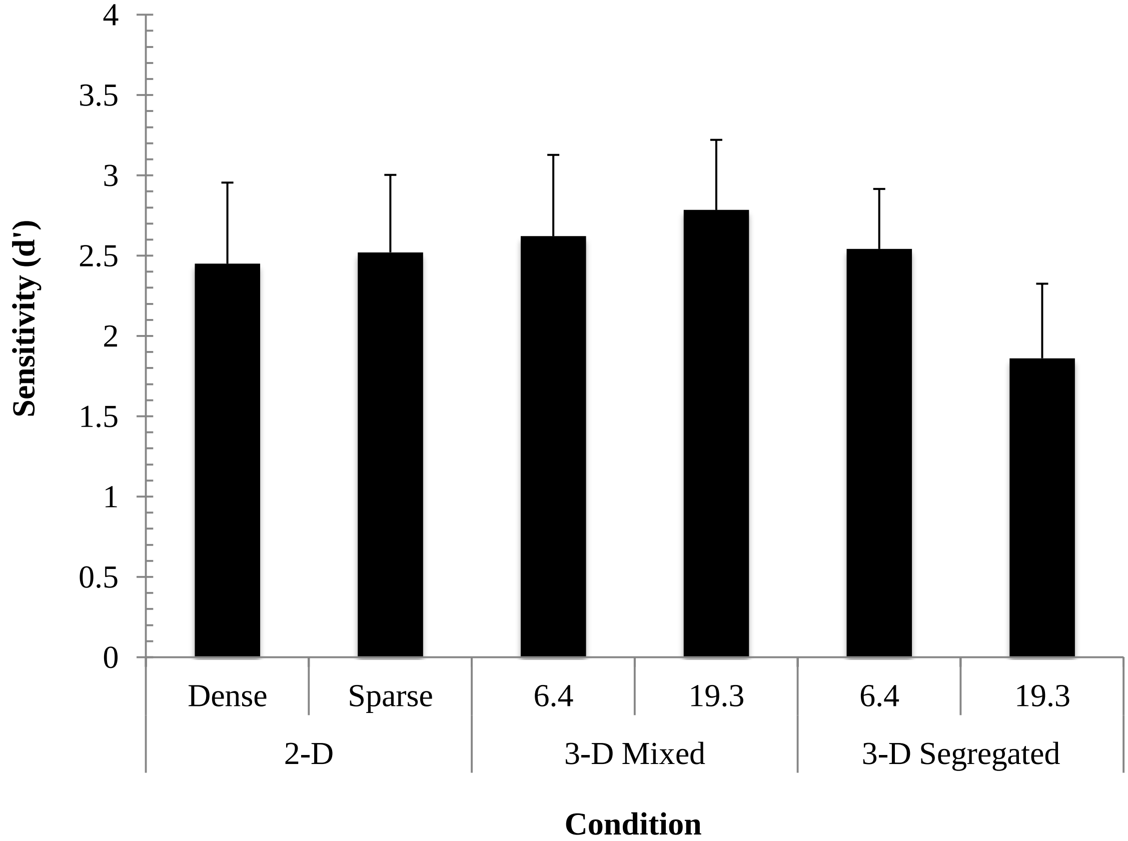

Figure 10. Object task d’ scores averaged across observers when disparity (arcmin) varied and the amount of time for encoding was increased (N = 10). Error bars represent one standard error of the mean.

CHAPTER V:

GENERAL DISCUSSION

This thesis examined whether viewing images in stereoscopic 3-D enhanced observers' ability to remember visual information stored in short-term memory. After many experiments, the consistent answer is no, visual short-term memory is not improved by viewing simple images in stereoscopic 3-D, compared to 2-D. This finding is consistent with Reeves and Lei (2014) that found dividing elements across disparity-defined planes does not affect VSTM, but conflicts with the conclusions drawn by Xu & Nakayama (2007), that found stereoscopic 3-D improved VSTM.

Summary

As the aim of this thesis was to determine the impact of stereoscopic 3-D on VSTM, it was critical to ensure that the experiments were evaluating only VSTM with no contribution from verbal memory. Therefore, a pilot experiment was performed to ensure that a preliminary verbal task was difficult enough to consume verbal short-term memory. The results indicated that *four* letters worked best, allowing for visual short-term memory to be used on a subsequent visual memory task. In all experiments, the 3-D conditions had two depth planes displayed simultaneously to ensure there was always relative depth present. Experiment 1 was conducted to assess both the influence of stereoscopic 3-D and the role of sample-test similarity on VSTM. Consistent with Awh et al. (2007), VSTM performance was better when there was a high similarity between the sample and test stimuli. As well, when the two depth planes were separated by a fixed amount of disparity (10.3 arcmin), VSTM was not influenced by the presence of depth or by the mode of segregation (mixed or segregated).

Experiment 2 was conducted to determine whether the lack of effect of 3-D on VSTM in Experiment 1 was tied to the specific fixed disparity used. In part a, the amount of disparity between depth planes was varied and the mode of segregation was interleaved across trials; in part b the amount of disparity was varied but conditions were blocked based on mode of segregation (mixed tested separately from segregated). Overall, VSTM performance was worse when disparity was varied (Experiment 2a) compared to when it was fixed (Experiment 1), and no difference was found in VSTM performance regardless of whether objects were presented in 2-D or 3-D. When disparity varied across trials, VSTM performance was better when the 3-D conditions were tested separately (Experiment 2b) than when they were interleaved (Experiment 2a), suggesting that observers use particular strategies based on the segregation of objects. For example, when objects are uniformly segregated across depth planes, they could group the items on each plane by similarity; when they are mixed across depth planes, they could group the items into a shape. Both encoding strategies become easier to employ when they can be used consistently across trials. This is consistent with Nakayama and Silverman's (1986) research on visual attention that found the presence of disparity-defined depth planes facilitated visual search. Despite this potential advantage, there was still no difference in VSTM performance between 2-D and 3-D conditions.

While the results of Experiment 2 showed no impact of stereoscopic 3-D on VSTM, varying disparity throughout a block may have hindered performance by increasing task difficulty. The lack of consistency of depth across trials could have eliminated the positive effects of a certain disparity. Experiment 3 evaluated various disparities in separate blocks, removing the possible confound of interleaving depth across trials. Varying the mode of segregation was detrimental when the amount of disparity varied (Experiment 2a), but did not negatively impact performance when disparity was fixed (i.e. Experiment 1 and Experiment 3). This likely relates to encoding strategies, which are most successful when there is consistency across trials. Should an observer use either a depth-related or grouping related encoding strategy it would be most effective when either depth or mode of segregation is held constant.

Although unlikely, it is possible that the larger disparities (19.3 arcmin and 25.8 arcmin) used in Experiments 2a, 2b, and 3, were seen as double (diplopic). Experiment 4 ruled out that possibility by showing that all disparities were seen as fused. Therefore, poor VSTM performance on the previous experiments could not be attributed to seeing the stimuli as diplopic.

Experiments 5 and 6 examined the role of non-depth specific encoding. When encoding was made more difficult, by adding distracting elements (Experiment 5), VSTM was degraded. When encoding was facilitated, by increasing the stimulus duration during encoding (Experiment 6), VSTM sensitivity was improved. These results, combined with the lack of impact of depth in these experiments, support the argument that the critical factor for influencing VSTM performance is encoding, rather than the presence of depth.

The role of encoding

As outlined above, differences in VSTM in this series of experiments are tied to encoding, and presence or absence of 3-D does not seem to affect this process. Encoding, the process of entering information into storage for later recall, is believed to take place over time, with features most closely related to the sensory input being encoded most rapidly (Shulman, 1970). There are large individual differences in the speed of encoding, particularly when it comes to visual information (Jannati, McDonald, & Di Lollo, 2015). This reflects the fact that VSTM encoding is influenced by many factors including: the duration of the retention interval, the symmetry and connectedness of elements being remembered, and the location of the to-be-remembered information on an object (Bankó & Vidnyánszky, 2010; Lai, Chien, Tai, & Hsu, 2009; Xu, 2002). While some information is intentionally stored through the use of active processes (e.g. rehearsing the information), other information can be unintentionally stored (Chen & Wyble, 2015). For instance, Chen and Wyble (2015) showed that when individuals are not instructed to remember particular information about a stimulus, the 2-D location, but not particular attributes (e.g. colour), are automatically encoded. This finding may apply to 3-D space, such that 3-D location is also automatically encoded. However, if this does occur, our results show that the information about 3-D layout is not exploited to improve VSTM.

Although the particular location of an object can be passively encoded (without overt attention), the encoding phase and capacity of visual short-term memory are typically linked to attention (Cowan, 2001; Xu & Chun, 2006). For example, it has been argued that attending to a stimulus can boost its apparent contrast, increasing its salience, which makes the stimulus easier to encode and remember (Carrasco, Ling & Read, 2004). It stands to reason that features that influence attention would also impact visual short-term memory; as outlined in the introduction researchers have proposed that stereoscopic depth could be one such feature. That is, Nakayama and Silverman (1986) found that the presence of depth-defined planes facilitates visual search (a task commonly used to assess attention). In their experiments, they found separation of elements across depth planes allows for parallel search strategies. Further research on depth and attention has shown that when perceptual load is low, attention can be reflexively allocated to locations in depth, regardless of whether depth is relevant to a task or not (Bauer, Plinge, Ehrenstein, Rinkenauer & Grosjean, 2012). More recently, Finlayson, Remington, Retell and Grove (2013) found that while the presence of depth can benefit attention, those benefits are tied to the task and stimulus configuration. In their experiments, the benefits of segmentation in depth, measured in terms of search efficiency, were only found when the target depth plane was known in advance. Moreover even when the target plane was cued before a trial, the benefit was not observed when the target and non-target plane required a conjunction search; a search in which the target and distractors share more than one similar visual property (e.g. colour, shape, orientation) (Treisman & Gelade, 1980). Finlayson et al. (2013) concluded that although segmentation of a search array into depth planes can facilitate visual search, stereopsis does not promote automatic preattentive segmentation in the way other elementary properties (e.g. colour) do (Finlayson et al., 2013). This finding is consistent with the results reported by Chau and Yeh (1995) who also found that depth can facilitate visual search, but depth does not improve visual search as strongly as the colour of a stimulus does. Therefore, depth is considered a weaker cue than other object features. Others have reported that when a target-relevant distractor is located on an unattended depth plane it interferes with visual search times, but this interference does not occur for target-irrelevant distractors (Theeuwes, Atchley & Kramer, 1998). The results of Theeuwes et al. (1998) provide compelling evidence that it is not always possible to inhibit information from an irrelevant depth plane.

Thus, while the presence of stereoscopic depth can facilitate attention-mediated visual search, this enhancement is specific to particular tasks and stimulus arrangements (e.g. Finlayson et al., 2013). Under other conditions, attention-mediated visual search is negatively influenced by the presence of distractor elements on neighbouring depth planes. In the experiments presented here the stimuli were clearly visible, and, apart from Experiment 5, there were no irrelevant distracting stimuli. Given this, and the work of Xu and Nakayama (2007) we anticipated that distribution of the elements in the display in two depth planes might make it easier for the visual system to encode the bracket orientations, particularly when each orientation occupied a different plane in depth. This was not the case.

It is possible that the impact of stereoscopic 3-D on VSTM is similar in some respects to its impact on visual attention. That is, under some circumstances, like those presented here, the segregation of information across multiple depth planes could provide an advantage, but observers do not reliably use this cue to aid recall. If depth information were somehow critical to the task, it is possible that observers could attend to a specific depth plane to assist encoding. In the experiments presented here, I followed the paradigm used by Xu and Nakayama (2007) who focussed an observers' ability to make use of 3-D to encode visual stimuli for short-term recall without explicit instructions to do so. Under these conditions, I have shown conclusively that the distribution of elements in depth does not aid VSTM.

Considerations

The object stimuli used throughout this thesis (oriented square brackets) are simple visual stimuli. Therefore, the findings in this thesis apply to simple visual objects. This may seem to be a limitation, however, if a benefit of 3-D for VSTM does exist, it would be most noticeable when there is less complexity. If the stimuli were more complex, a number of visual attributes could contribute and potentially interact to influence VSTM making it difficult to isolate any single factor.

As mentioned above, to be consistent with the methodology used in prior studies on this topic (Reeves & Lei, 2014; Xu & Nakayama, 2007), the stereoscopic depth was not a property necessary to completing the visual memory task. Rather, depth was provided as an optional source of information, which could have been used to facilitate encoding. A different experimental design would have to be used to assess whether or not depth from disparity is directly encoded in VSTM. If retaining depth information were necessary for completing a task, it seems likely that observers would be able to use such information; however additional experiments that are outside of the scope of this thesis would be needed to test this proposal.

Conclusions

In conclusion, visual short-term memory for simple visual objects is not improved by organizing objects into separate stereoscopic planes. This finding conflicts with Xu and Nakayama (2007), however they sequentially (rather than simultaneously) presented their stimuli, which effectively eliminated stereoscopic depth. On the other hand, these results are consistent with Reeves and Lei (2014), who used simultaneous presentation of stereoscopic depth planes.

While I used simple line stimuli in my experiments the results may generalize to more complex 3-D stimuli, like those used in advertising. If so, I would argue that the use of stereoscopic 3-D to display imagery on multiple depth planes would provide little recall advantage in the short term. It is possible however; that a longer-term impact on recall could be obtained if a 3-D object (product) was placed in an otherwise 2-D series of images. Note that in this scenario the critical property is the enhanced salience due to the novelty of the 3-D, not the distribution of elements in depth per se.

While stereopsis is a powerful depth cue, it is rarely present in isolation in a natural setting; rather it is typically present with other depth cues (e.g. shadows) and stimulus properties (e.g. colour). Individuals therefore have less experience using stereopsis independently to facilitate encoding compared to other features that are often viewed in isolation (e.g. colour). Because colour is a salient property that is sometimes viewed on its own, it can more easily be retained in conjunction with other properties such as orientation (Luck & Vogel, 1997). Overall, depth from stereopsis in isolation does not seem to be sufficiently salient to be used to facilitate visual short-term memory.

REFERENCES

Agranoff, B., & Davis, R., & Brink, J. J. (1965). Chemical studies on memory fixation in goldfish*. *Brain Research*, *1*(5), 303–309.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622–628. http://doi.org/10.1111/j.1467-9280.2007.01949.x

Bankó, E. M., & Vidnyánszky, Z. (2010). Retention interval affects visual short-term memory encoding. *Journal of Neurophysiology*, *103*(3), 1425–1430. http://doi.org/10.1152/jn.00868.2009

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

Baddeley, A. D. (1992). Working memory. *Science, 255*(5044), 556-559. http://doi.org/10.1016/j.cub.2009.12.014

Baddeley, A.D., & Hitch, G. (1974). *Working memory*. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory. 8,* 47-89. New York: Academic Press.

Bauer, D., Plinge, A., Ehrenstein, W. H., Rinkenauer, G., & Grosjean, M. (2012). Spatial orienting of attention in stereo depth. *Psychological Research-Psychologische Forschung*, *76*(6), 730–735. http://doi.org/10.1007/s00426-011-0396-6

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B, 57*, 289-300.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(38), 14325–14329. http://doi.org/10.1073/pnas.0803390105

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433-436.

Brown, J. (1958). Some Tests of the Decay Theory of Immediate Memory. *Quarterly Journal of Experimental Psychology, 10,* 12-21

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*(3), 308–313. http://doi.org/10.1167/12.9.1387

Chau, a W., & Yeh, Y. Y. (1995). Segregation by color and stereoscopic depth in three-dimensional visual space. *Perception & Psychophysics*, *57*(7), 1032–1044. http://doi.org/10.3758/BF03205462

Chen, H., & Wyble, B. (2015). The location but not the attributes of visual cues are automatically encoded into working memory. *Vision Research*, *107*, 76–85. http://doi.org/10.1016/j.visres.2014.11.010

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*(1), 87–114; discussion 114–185. http://doi.org/10.1017/S0140525X01003922

Cutting, J. E. & Vishton, P. M. (1995). Perceiving layout and knowing distances: The interaction, relative potency, and contextual use of different information about depth. In W. Epstein & S. Rogers (Eds.) *Perception of space and motion* (pp. 96-117). San Diego, CA: Academic Press.

Druzgal, T. J., & D'Esposito, M. (2001). Activity in fusiform face area modulated as a function of working memory load. *Cognitive Brain Research*, *10*(3), 355–364. http://doi.org/10.1016/S0926-6410(00)00056-2

Farah, M. J. (1989). Mechanisms of Imagery-Perception Interaction. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(2), 203–211.

Finlayson, N. J., Remington, R. W., Retell, J. D., & Grove, P. M. (2013). Segmentation by depth does not always facilitate visual search. *Journal of Vision*, *13*(2013), 1–14. http://doi.org/10.1167/13.8.11.doi

Goldstein, E. B. (2010*). Sensation and Perception* (8th edition). New York: Brooks Cole Publishing.

Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics. New York: Wiley

Hales, J. B., & Brewer, J. B. (2012). The Path to Memory Is Guided by Strategy: Distinct Networks Are Engaged in Associative Encoding under Visual and Verbal Strategy and Influence Memory Performance in Healthy and Impaired Individuals. *Journal of Cognitive Neuroscience*, *24*(6), 1398–1410. http://doi.org/10.1162/jocn_a_00220

Howard, I. P., & Rogers, B. J. (2012). Perceiving in Depth, Vol. 2, Stereoscopic Vision. New York: Oxford University Press.

Hulme, C., Maughan, S., & Brown, G. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span., *701*, 685–701. http://doi.org/10.1016/0749-596X(91)90032-F

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G., Martin, S., & Stuart, G. (1997). Word frequency effects on short-term memory tasks: Evidence for a redintegration proces in immediate serial recall., *23*(5). Retrieved from http://discovery.ucl.ac.uk/1327607/

Ihssen, N., Linden, D. E. J., & Shapiro, K. L. (2010). Improving visual short-term memory by sequencing the stimulus array. *Psychonomic Bulletin & Review*, *17*(5), 680–686.

Jannati, A., Lollo, V. Di, & Mcdonald, J. J. (2015). Individual differences in rate of encoding predict estimates of visual working memory capacity (K), *69*(2), 778–782.

Jefferies, E., Frankish, C. R., & Lambon Ralph, M. A. (2006). Lexical and semantic binding in verbal short-term memory. *Journal of Memory and Language*, *54*(1), 81–98. http://doi.org/10.1016/j.jml.2005.08.001

Keogh, R., & Pearson, J. (2011). Mental imagery and visual working memory. *PLoS ONE*, *6*(12). http://doi.org/10.1371/journal.pone.0029221

Kim, H. (2011). Neural activity that predicts subsequent memory and forgetting: A meta-analysis of 74 fMRI studies. *NeuroImage*, *54*(3), 2446–2461. http://doi.org/10.1016/j.neuroimage.2010.09.045

Konstantinou, N., Bahrami, B., Rees, G., & Lavie, N. (2012). Visual Short-term Memory Load Reduces Retinotopic Cortex Response to Contrast. *Journal of Cognitive Neuroscience*, *24*(11), 2199–2210. http://doi.org/10.1162/jocn_a_00279

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., … Alpert, N. M. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science (New York, N.Y.)*, *284*(5411), 167–170. http://doi.org/10.1126/science.284.5411.167

Lai, H. C., Chien, S. L., Tai, S.K., Hsu, H. Y. (2009). "Visual short-term memory for abstract patterns across encoding durations: effects of symmetry, element connectedness, and probe quadrant" *Perception* **38** ECVP Abstract Supplement, p. 178

Luck, S. J. (2007). Visual short-term memory. *Scholarpedia*, *2*(6): 3328.

Luck, S. J., & Hollingworth, A. (2008). *Visual Memory*. New York: Oxford University Press.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. http://doi.org/10.1038/36846

Macmillan, N. A., & Creelman, C. D. (2005). Detection Theory: A User's Guide. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Nakayama, K., & Silverman, G. H. (1986). Serial and Parallel Processing of Visual Feature Conjunctions. *Nature, 320*(6059), 264–265. http://doi.org/10.1038/320264a0

Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature*, *362*(6418), 342–345. http://doi.org/10.1038/362342a0

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into 632 movies. *Spatial Vision*, *10*(4), 437-442

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *The Behavioral and Brain Sciences*, *31*(2), 109–130; discussion 130–178. http://doi.org/10.1017/S0140525X08003543

Perky, C. W. (1910). An Experimental Study of Imagination. *The American Journal of Psychology*, *21*(3), 422-452.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*(2), 283–290.

Raymond, J. E., Shapiro, K. L., Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *Journal of Experimental Psychology: Human perception and Performance,* 18(3), 849-60.

Reeves, A., & Lei, Q. (2014). Is visual short-term memory depthful? *Vision Research*, *96*, 106–112. http://doi.org/10.1016/j.visres.2014.01.010

Ruchkin, D. S., Grafman, J., Cameron, K., & Berndt, R. S. (2003). Working memory retention systems: A state of activated long-term memory. *Behavioral and Brain Sciences*, *26*(6), 709–777.

Schraagen, J. M., Chipman, S. F., & Shalin, V. L. (2000). *Cognitive Task Analysis*. Taylor & Francis.

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science : A Journal of the American Psychological Society / APS*, *20*(2), 207–214. http://doi.org/10.1111/j.1467-9280.2009.02276.x

Shallice, T., & Warrington, E. K. (1970). Independent Functioning of Verbal Memory Stores: a Neuropsychological Study. *Quarterly Journal of Experimental Psychology*, *22:2*, 261–273. http://doi.org/10.1080/00335557043000203

Shulman, H. G. (1970). Encoding and retention of semantic and phonemic information in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 499–508. http://doi.org/10.1016/S0022-5371(70)80093-7

Simons, D. J., & Rensink, R. a. (2005). Change blindness : Past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16–20.

Soto, D., Wriglesworth, A., Bahrami-Balani, A., & Humphreys, G. W. (2010). Working memory enhances visual perception: evidence from signal detection analysis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(2), 441–456. http://doi.org/10.1037/a0018686

Standing, L. (1973). Learning 10,000 pictures. *The Quarterly Journal of Experimental Psychology*, *25*(2), 207–222. http://doi.org/10.1080/14640747308400340

Theeuwes, J., Atchley, P., & Kramer, A. F. (1998). Attentional control within 3-D space. *Journal of Experimental Psychology-Human Perception and Performance*, *24*(5), 1476–1485. http://doi.org/10.1037//0096-1523.24.5.1476

Treisman, A. (1988). Features and objects: the fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *40*(2), 201–237. http://doi.org/10.1080/02724988843000104

Treisman, A. M., & Gelade, G. (1980). Feature-Integration Theory of Attention. *Cognitive Psychology*, *12*(1), 97–136. http://doi.org/10.1016/0010-0285(80)90005-5

Van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, *109*(22), 8780–8785. http://doi.org/10.1073/pnas.1117465109

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology-Human Perception and Performance*, *32*(6), 1436–1451. http://doi.org/10.1037/0096-1523.32.6.1436

Xu, Y. (2002). Encoding color and shape from different parts of an object in visual short-term memory. *Perception & Psychophysics, 64*(8), 1260–1280. http://doi.org/10.3758/BF03194770

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*(7080), 91–95. http://doi.org/10.1038/nature04262

Xu, Y., & Nakayama, K. (2007). Visual short-term memory benefit for objects on different 3-D surfaces. *Journal of Experimental Psychology: General*, *136*(4), 653–662. http://doi.org/http://dx.doi.org/10.1037/0096-3445.136.4.653

APPENDIX A:
Proportion Correct Graphs for Experiments 1, 2a, 2b, 3, 5 and 6

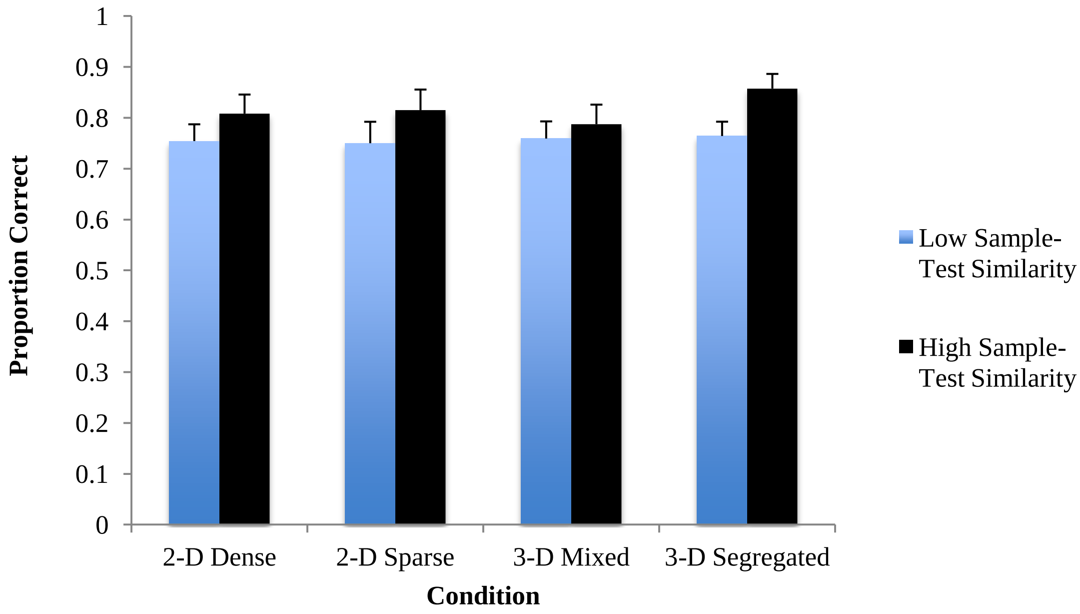

Figure A1. Experiment 1 object task proportion correct scores averaged across observers for the low sample-test similarity (N = 16) and high sample-test similarity (N = 15) experiment. Error bars represent one standard error of the mean.

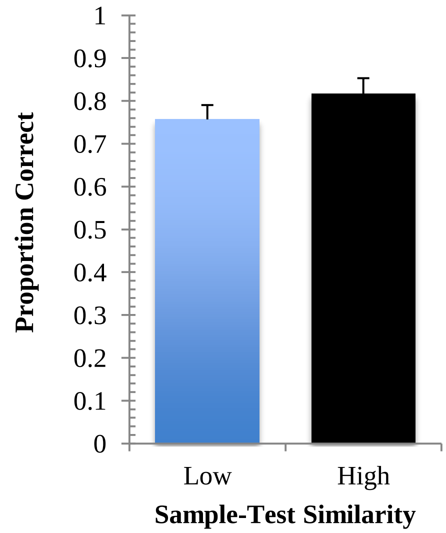

Figure A2. Experiment 1 object task proportion correct scores averaged for observers that took part in both the low and high sample-test similarity experiments (N = 12). Error bars represent one standard error of the mean.

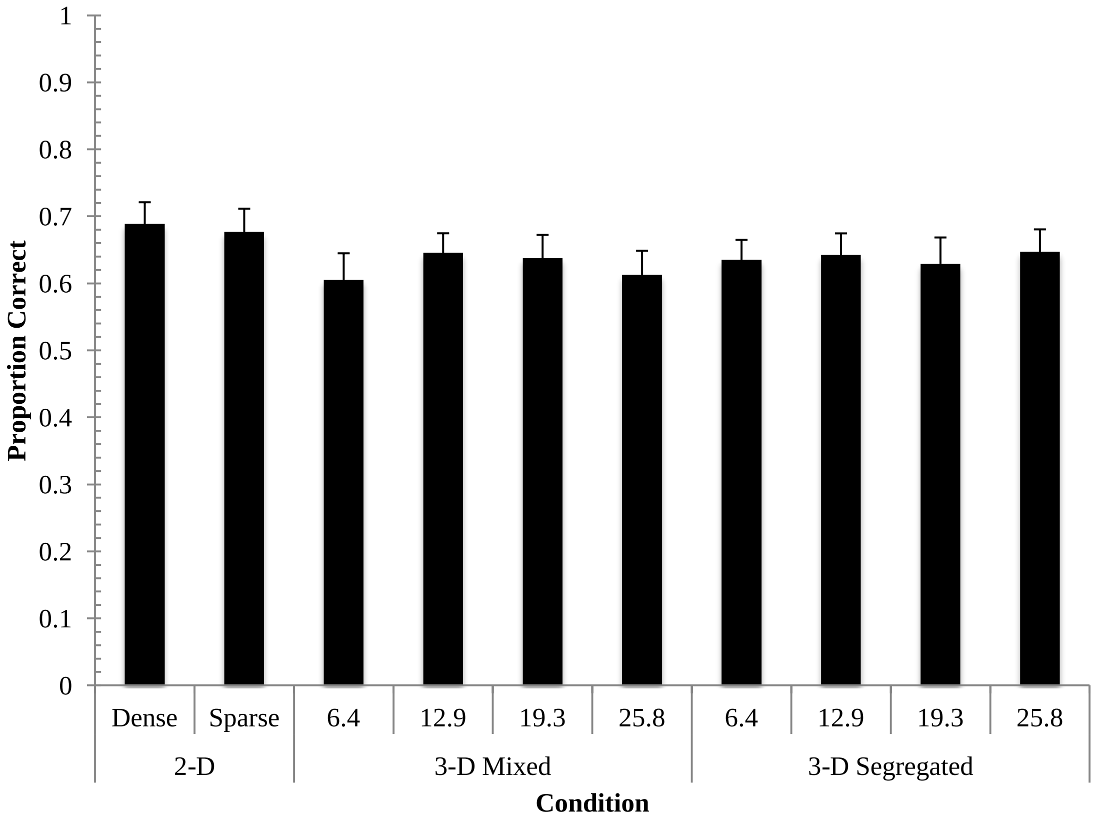

Figure A3. Experiment 2a object task proportion correct scores averaged across observers when disparity (arcmin) varied and 3-D conditions were interleaved (N = 21). Error bars represent one standard error of the mean.

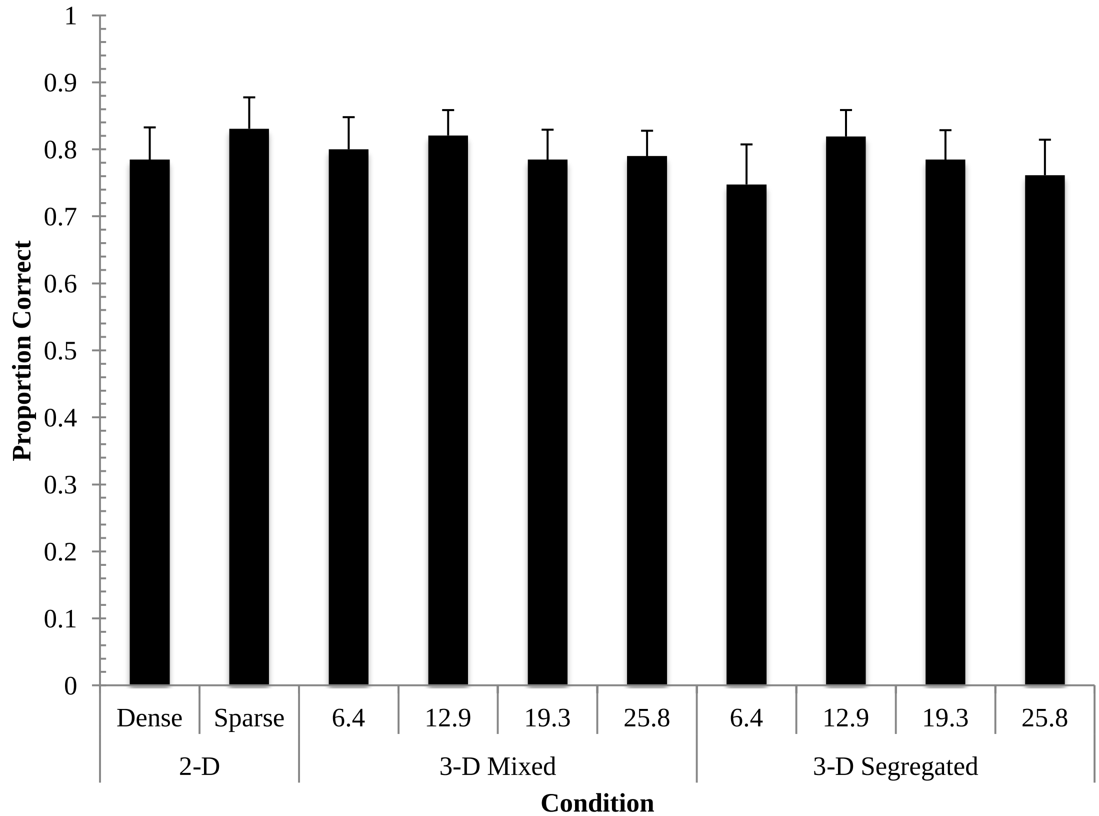

Figure A4. Experiment 2b object task proportion correct scores averaged across observers when disparity (arcmin) varied and 3-D conditions were blocked (N = 12). Error bars represent one standard error of the mean.

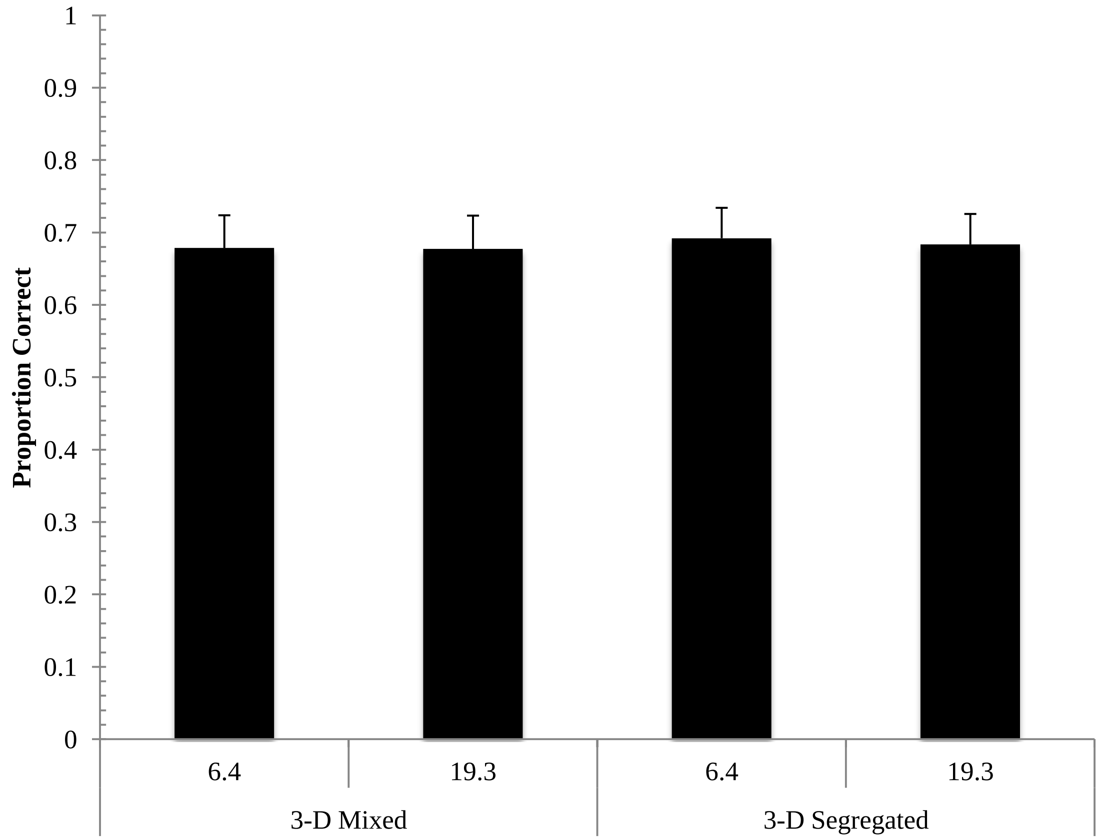

Figure A5. Experiment 3 object task proportion correct scores averaged across observers when 3-D conditions were interleaved but were blocked based on disparity (arcmin) (N = 14). Error bars represent one standard error of the mean.

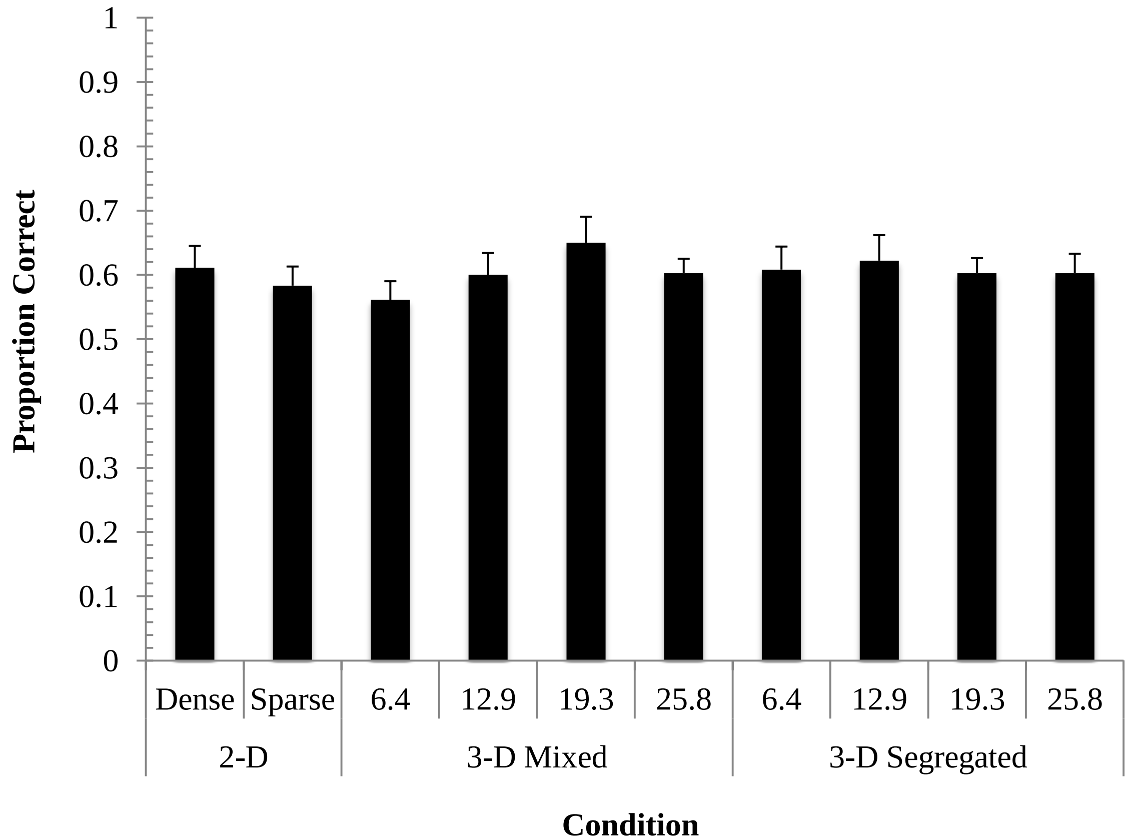

Figure A6. Experiment 5 object task proportion correct scores averaged across observers when disparity (arcmin) varied and distracting elements were present (N = 18). Error bars represent one standard error of the mean.

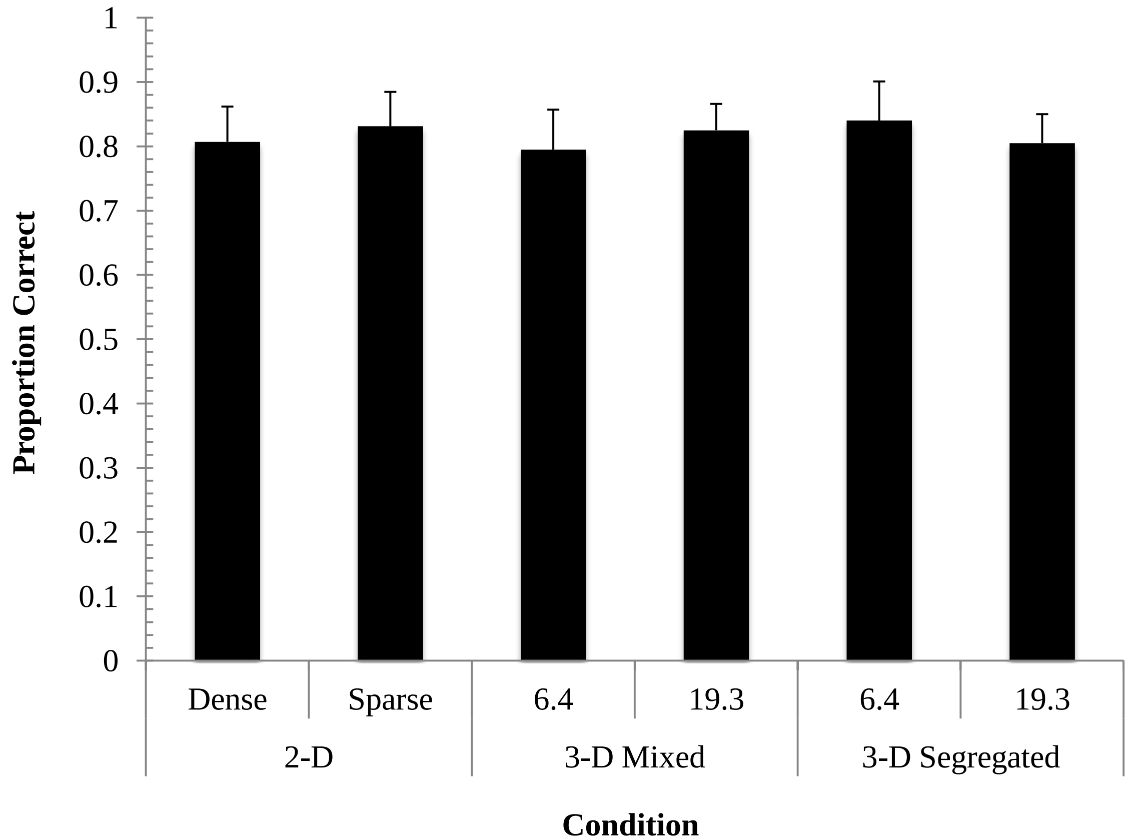

Figure A7. Experiment 6 object task d’ scores averaged across observers when disparity (arcmin) varied and the amount of time for encoding was increased (N = 10). Error bars represent one standard error of the mean.