8-2019

Identifying key markers of visual short-term memory: How looking dynamics and physiology can inform our interpretation of change detection

Bret Eschman
University of Tennessee, beschma1@vols.utk.edu

Recommended Citation

Eschman, Bret, "Identifying key markers of visual short-term memory: How looking dynamics and physiology can inform our interpretation of change detection. " PhD diss., University of Tennessee, 2019.
https://trace.tennessee.edu/utk_graddiss/5608

To the Graduate Council:

I am submitting herewith a dissertation written by Bret Eschman entitled "Identifying key markers of visual short-term memory: How looking dynamics and physiology can inform our interpretation of change detection." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Psychology.

Shannon Ross-Sheehy, Major Professor

We have read this dissertation and recommend its acceptance:

Aaron Buss, Gordon Burghardt, Jessica Hay, Greg Reynolds

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Identifying key markers of visual short-term memory: How looking dynamics and physiology can inform our interpretation of change detection

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Bret Thomas Eschman
August 2019

ACKNOWLEDGEMENTS

I would like to thank both of my parents for their constant support, patience, and encouragement throughout all of my academic endeavors. I would like to thank my wife, Ruth, for her tireless contribution to both my physical and emotional well-being during this process. Thank you to all of the past and present members of the Infant Visual Cognition Lab for all of the help on all of my projects over the years—I could not have done it without you all! I would also like to thank my committee members, Aaron Buss, Gordon Burghardt, Jess Hay, and Greg Reynolds, for their patience and their valued input on this project. It takes a village to raise a graduate student, and I am fortunate enough to be the product of an amazing community of researchers. Finally, I would like to thank my mentor, Shannon Ross-Sheehy for everything that she has done for me over the past 5 years. The passion with which you approach your research and mentorship has made a lasting impact on my career. I can only hope to make that same impact on students in the future.

ABSTRACT

Visual short-term memory (STM) is a foundational component of general cognitive ability that develops rapidly during the first year of life. Currently, it is unknow if visual STM performance in infancy reflects a similar memory mechanism used by adults. This is due to significant differences in the tasks used to measure visual STM performance in infant and adults. The current project has identified key behavioral and physiological indexes of visual STM performance in infants by utilizing data collected from adult participants in a similar task. In Experiment 1, adult visual dynamics were assessed during a change-detection task, and several key behaviors identified. In Experiment 2, these behaviors were subsequently observed in infants and adults while performing a similar change-detection task. Experiment 3 then applied infant-specific adaptations to an adult change-detection procedure, and again, found significant similar patterns of responding. Experiment 4 proposed a novel visual STM assessment technique, shedding light on the extent to which infant performance is uniquely influenced by incidental attention to individual array items. Results demonstrated that the order of fixation affected subsequent performance on a change-detection task. Combined, these results have identified an informative metric for understanding change detection in both infant and adult populations and have provided researchers with a novel method of measuring a cornerstone of cognitive development, visual STM. Taken together, results from these tasks demonstrate that visual dynamics such as saccade count, run count, average fixation duration, and changes in pupil size may be an ideal means of assessing visual STM ability in both infants and adults.

Keywords: *Visual short-term memory, Change detection, Visual dynamics, Task evoked pupillary responses*

PREFACE

“There is probably no part of the human body other than the human eye where I feel so intuitively that we have access to the innermost workings of the mind…”

-Eckhard H. Hess (1975)

TABLE OF CONTENTS

CHAPTER ONE INTRODUCTION AND BACKGROUND 1

Adult Visual Short-Term and Working Memory: Techniques and Measurement . 7

Physiological Measures of STM: Task-Evoked Pupillary Responses 10

Infant Visual Short-Term Memory: Techniques and Measurement 12

Current Project 25

CHAPTER TWO EXPERIMENT 1: Assessing the Behavioral and Physiological Correlates of STM Maintenance in Adults 27

Method 29
Participants 29
Apparatus 30
Stimuli 30
Design and Procedure 31

Results 33
Eye Tracker Event Parsing 33
Percent Correct 33
Saccade Count 34
Run Count 37
Average Fixation Duration 39
Capacity Estimates 41
Pupil Analyses 43

Discussion 46

CHAPTER THREE Infant and Adult Passive Change-Detection Task 48

EXPERIMENT 2A: Do Infants Demonstrate Similar Memory-Dependent Visual Dynamics as Adults? 49

Method 49
Participants 49
Apparatus 49
Stimuli 50
Design and Procedure 53

Results 53
Saccade Count 53
Run Count 54
Average Fixation Duration 57
Change Preference 57
Pupil Analyses 59

Discussion 62

EXPERIMENT 2B: Do Adults Demonstrate Similar Memory-Dependent Visual Dynamics in a Passive Change-Detection Task? 64

Method 64

Participants 64
Stimuli and Apparatus 64

Results 65
Saccade Count 65
Run Count 65
Average Fixation Duration 68
Pupil Analyses 68

Discussion 73

CHAPTER FOUR EXPERIMENT 3: Do Infant Specific Adaptations in Change-Detection Tasks Influence Adult Visual Behavior? 75

Method 76
Participants 76
Stimuli and Apparatus 77
Design and Procedure 79

Results 79
Percent Correct 79
Saccade Count 80
Run Count 84
Average Fixation Duration 84
Capacity Estimates 87
Pupil Analyses 89
Encoding Period Analyses 89

Discussion 91

CHAPTER FIVE EXPERIMENT 4: A New Approach to Visual STM Assessment 93

Method 94
Participants 94
Stimuli and Apparatus 95
Design and Procedure 95

Results 96
Saccade Count 96
Run Count 96
Average Fixation Duration 98
Change Preference 100
Pupil Analyses 102

Discussion 104

CHAPTER SIX GENERAL DISCUSSION AND CONCLUSIONS 105

Is Change Preference Appropriate for Estimating Visual STM Capacity? 110

Can You Use an *N*-Back Approach to Estimate Visual STM Capacity? 112

Are Passive Change-Detection Tasks Fundamentally Different than those that Have an Overt Behavioral Response? 113

Using Visual Dynamics in the Future 114

General Conclusions 116

LIST OF REFERENCES 117
APPENDIX 123
Vita 125

LIST OF TABLES

Table 1: Analysis of Variance Visual Dynamics Results (Experiment 1) 36
Table 2: Capacity Estimate and Visual Dynamic Correlations (Experiment 1) 42
Table 3: Analysis of Variance Visual Dynamics Results (Experiment 2A) 55
Table 4: Analysis of Variance Visual Dynamics Results (Experiment 2B) 66
Table 5: Analysis of Variance Visual Dynamics Results (Experiment 3) 82
Table 6: Capacity Estimate and Visual Dynamic Correlations (Experiment 3) 88
Table 7: Analysis of Variance Visual Dynamics Results (Experiment 4) 97
Table 8: Change Preference and Visual Dynamics (Experiment 4) 101
Table 9: Summary Table of All Significant Differences in Visual Dynamics 106
Table 10: Experiment 1-4 RGB and Luminance Values 124

LIST OF FIGURES

Figure 1: Baddeley and Hitch (1974) Model of Working Memory 4
Figure 2: Schematic Representation of a Trial Adapted from Ross-Sheehy et al., 2003 15
Figure 3: Schematic Representation of a Trial Adopted from Oakes et al., 2013 19
Figure 4: Schematic Representation of a Trial Used in Ross-Sheehy & Eschman, 2019 21
Figure 5: Change Preference Minus Chance for Each Set Size, as a Function of Age and Set Size (Ross-Sheehy & Eschman, 2019) 22
Figure 6. Change Preference as a Function of Prior Fixations to the To-Be-Changed Item During the Sample Array 24
Figure 7: Experiment 1, Change Detection Task (Luck & Vogel, 1997), Set Size 8 Change Trial 28
Figure 8: Average Percent Correct as a Function of Change Status and Set Size. 35
Figure 9: Saccade Count as a Function of Response Accuracy and Change Status During the Test Array 38
Figure 10: Average Fixation Duration as a Function of Response Accuracy 40
Figure 11: TEPRs as a Function of Response Accuracy and Bin. 45
Figure 12: Experiment 2, Change Detection Task Set Size 3 Change Trial 51
Figure 13: Infant Run Count as a Function of Age, Change Status, and Set Size 56
Figure 14: Infant TEPRs as a Function of Bin and Set Size 61
Figure 15: Infant TEPRs as a Function of Bin and Change Status 61
Figure 16: Adult Run Count as a Function of Change Status and Set Size 67
Figure 17: Adult Average Fixation Duration as a Function of Change Status and Set Size 69
Figure 18: Adult TEPRs as a Function of Bin and Change Status 71
Figure 19: Adult TEPRs as a Function of Bin and Set Size 72
Figure 20: Experiment 3, Change Detection Task Set Size 6 Change Trial 78
Figure 21: Average Percent Correct as a Function of Change Status and Set Size 81
Figure 22: Saccade Count During the Test Array as a Function of Response Accuracy 83
Figure 23: Run Count During the Test Array as a Function of Response Accuracy 85
Figure 24: Average Fixation Duration as a Function of Response Accuracy 86
Figure 25: TEPRs as a Function of Response Accuracy and Bin 90
Figure 26: Run Count as a Function of Period, Age, and *N*-Back Value 99
Figure 27: Change Preference and a Function of Age and *N*-Back Value 101
Figure 28: TEPRs as a Function of Age and *N*-Back Value 103

CHAPTER ONE

INTRODUCTION AND BACKGROUND

As we look around our environment, eye movements and blinks break the flow of visual information into multiple discrete “snapshots”. In order to make sense of these disjointed snapshots, visual short-term memory (STM) may serve to bind this visual input into a coherent visual percept. Specifically, visual STM allows individuals to rapidly store a subset of visual information and then to access that information quickly, in order to act on a goal (Luck, 2007). For example, while driving a car down a busy interstate, we must be able to rapidly encode and remember the locations of the other cars, so we can act on that information efficiently—such as abruptly switching lanes if a hazard is detected. Importantly, the number of items that one can keep active in visual STM is highly capacity limited. This incoming visual information is readily overridden and susceptible to decay, drift, or termination over time (Luck, 2007; Phillips, 1974; Shaffer & Shiffrin, 1972). If we do not refresh this visual information (i.e. continue to check our surroundings), or replace it with additional visual stimuli, we will lose the relevant information (location of the cars).

Visual STM is important and plays a critical role in many of our foundational cognitive abilities. For example, it acts as a buffer or temporary storage while processing naturally occurring stimuli (i.e. the previous location of the cars while scanning the road for additional information). Visual STM is more durable than iconic memory (visual sensory memory that precedes visual STM), and persists through eye movements, blinks, and other visual interruptions (Luck, 2007). In addition, early in development, visual STM may allow for the integration of low-level feature information such as color, with higher-level cognitions such as object recognition (Bettencourt & Xu, 2016), as well as comparison and categorization (Kwon, Luck, & Oakes, 2013).

Given that visual STM is a foundational cognitive capability, it is important that we understand early markers of visual STM development in infants, and how limits in visual STM may be related to limits in general cognitive functioning. Currently, there is little work connecting infant and adult measures of visual STM; thus, it is not clear if infant visual STM is related to later memory performance, or if infant and adult tasks are tapping into the same memory systems. The reason for this lack of research connectivity stems from the lack of visual STM testing techniques that are suitable for use in both infants and adults.

In adults, visual STM capacity is typically measured using a delayed match-to-sample, or "change-detection" task. Most often, adults are shown an array of stimuli (i.e., *sample*), followed by a brief retention interval, and finally the presentation of a second array (i.e., *test*) that is either identical to the first array (i.e., "same") or different in some way (i.e., "different"). In tasks such as this, STM is likely a *working memory* (WM) system, both because adults are explicitly attempting to remember something, and because the memory representation is subsequently used to make a decision (see Luck, 2008 for a review). Specifically, visual WM allows information to be stored in a place that can be quickly accessed in service of a cognitive task (Baddeley & Hitch, 1974). In these tasks, researchers are directly assessing a specific component of Baddeley's model of WM, the visuospatial sketchpad (Baddeley & Hitch, 1974; Figure 1). In this, and later updates of this model (Baddeley & Logie, 1999; Baddeley & Hitch, 1994), the visuospatial sketchpad is a short-term cache for storing visual information. Implications of this conception are that limits in the size of the cache necessarily limit visual WM ability, even if processing is otherwise adequate. Thus, measuring the capacity of this visual cache may help us understand what limits visual STM performance on typical lab-based tasks.

Historically, infant researchers have opted for the term STM rather than WM. This is due to the fact that theories of WM typically include both storage components and components for active memory manipulation (Baddeley & Hitch, 1974). Because prelinguistic infant participants cannot be explicitly instructed to perform a given task, we cannot be certain they are using their STM in service of the task (i.e. a WM system). This lack of explicit instruction is one key way in which infant and adult tasks may differ, making comparisons across these populations difficult, if not impossible.

In addition to implicit/explicit instruction differences, infant tasks are usually *qualitatively different* than adult tasks. For example, infant visual STM, has often been assessed using a *preferential looking* task, which is considerably different from the adult visual WM measures mentioned above. In these tasks, infants are given the choice to look at one of two displays (one containing an

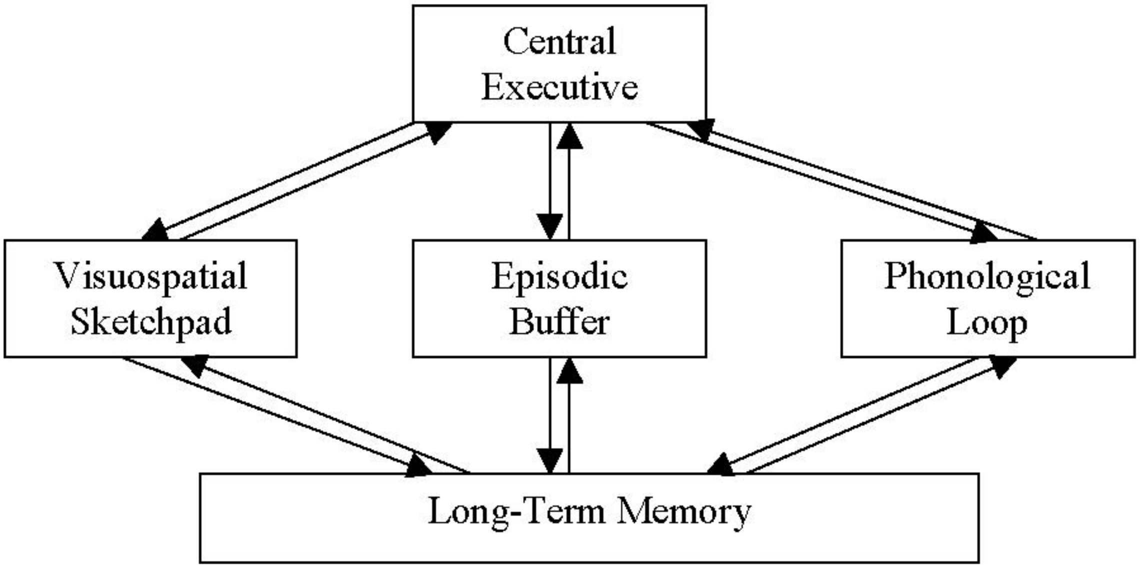


Figure 1: Baddeley and Hitch (1974) Model of Working Memory

array of blinking, color-changing squares, and another array containing blinking, non-changing squares).

Change preference, or the proportion of time that infants spend looking to the changing array relative to total looking across both arrays, is often used as an index of change detection and STM (Oakes, Messenger, Ross-Sheehy, & Luck, 2009; Ross-sheehy, Oakes, & Luck, 2003, 2011). Although these procedures have revealed capacity-like effects that increase with development (Oakes, Messenger, Ross-Sheehy, & Luck, 2009; Ross-Sheehy, Oakes, & Luck, 2003, 2011), it is very likely that that the demands of the two different tasks may result in performance differences. Recently, in a review of WM capacity, Simmering (2016) posits that, “it is possible that the same processes operate across all tasks, [but] the demands are minimized in simple tasks, reducing the ability to detect correlations.” As a result, simple tasks that are related to cognitive skills like STM or WM, can produce performance differences if the demands of said task are low. For example, Unsworth and Engle (2007) demonstrated that when a task was simple/easy, it was correlated less with higher cognitive skills than when the task was more difficult. Specifically, when a simple span task was scaled to beyond an individual’s capacity limit, correlations between performance and general intelligence were strengthened. Hence, while performance between these two tasks may differ, it is still likely the same underlying process is being used.

While researchers have made significant progress in recent years, it is currently unclear if *visual* STM performance as measured via change-detection tasks, is related to either concurrent or later cognitive performance. There has however, been a considerable amount of work with children that demonstrates a clear link between *general* WM and/or STM ability and IQ (Alloway et al., 2009; Alloway & Passolunghi, 2011), math achievement (Bull, Espy, & Wiebe, 2008) and general fluid intelligence (Kane & Engle, 2002). For example, performance on a series of STM tasks was found to be the best predictor of number ranking and arithmetic skills in 8-year-old children (Alloway & Passolunghi, 2011). Similarly, STM performance at age 4 using the Corsi Block Task predicted math ability at 7 years of age (Bull et al., 2008). In addition to predicting general math ability, STM has also been shown to be related to both reasoning and verbal comprehension tasks. For example, children with a low STM capacity are more likely to perform poorly on a number of key learning measures (i.e. reading and math) and are more likely to have attentional difficulties (i.e. more distractible), leading to careless mistakes (Alloway et al., 2009). Because many of these tasks rely on behavioral approaches for estimating STM capacity (e.g., card sorting tasks), the extent to which *visual* STM plays a role in these general cognitive outcomes is unknown. In order to assess this in the future, researchers need a method of isolating visual STM and more importantly, a way to measure it in a continuous manner. The series of tasks that are outlined in this dissertation will provide an approach that accomplishes both of these needs.

Adult Visual Short-Term and Working Memory: Techniques and Measurement

As mentioned above, most adult visual STM work is based on some form of change detection. In their seminal study, Luck and Vogel (1997) utilized a change-detection task to determine how many items typical adults could hold in visual STM (i.e. their visual STM capacity). Subjects viewed 1-12 colored squares for 100ms, followed by a 900ms retention interval, and finally the presentation (for 2000ms) of either that exact same array, or the same array with one color changed. On each trial, participants were instructed to press a button to indicate if the arrays were the same or different. Accuracy at each set size was used to estimate each participant's STM capacity. Results demonstrated that participants were nearly perfect for set sizes 1-3, with a systematic decline in performance from set size 4-12. Luck and Vogel (1997) concluded that on average, adults have a visual STM capacity of around 3-4 items. Though there are multiple approaches for estimating capacity based on response accuracy in these types of tasks (Cowan, 2001; Pashler, 1988; d-Prime), they are all generally based on the assumption that participants are typically correct for arrays *within* their capacity. For example, a participant with a STM/WM capacity of around 3 items may be perfect for set sizes from 1-3, but performance on set size 4 will suffer, possibly resulting in a guess. These guesses should be systematically more likely as set size increases beyond capacity. Although imperfect, these simple capacity calculations have motived a great deal of work aimed at understanding individual differences in STM/WM. For example, WM capacity scores positively correlate with reading comprehension, complex learning, and reasoning (Daneman & Carpenter, 2004), as well as general intelligence (Engle, 2002).

However, individual differences may not *all* be driven by differences in STM/WM storage, *per se*. Recent work suggests lapses in attention (Unsworth & Robison, 2015; Vogel, McCollough, & Machizawa, 2005) may also contribute to individual differences in STM performance, and could be an important predictor of general STM ability. For example, performance has been shown to suffer when pupillary indexes of attentional vigilance signal a subject may be “off task” (Unsworth & Robison, 2015, 2016). In another study, STM performance and subsequent capacity estimates were shown to be highly influenced by selective attention—participants who were relatively good at filtering out task-irrelevant information demonstrated particularly high STM capacity (Vogel et al., 2005). Finally, in their review, Astle and Scerif (2011) refer to items in visual STM as having a “privileged state of activation.” More specifically, they contend that because there is so much potential information competing for access into visual STM, the ability to decide what gets in and what does not (i.e. attentional filtering) must be tightly linked to measures of visual STM capacity. While capacity estimates that include these additional attentional indicators may increase accuracy in capacity estimation, it is important to note that attention is only part of the story.

In addition to thinking about the potential role attention has on measuring visual STM and WM performance, it is also important to think about some of the previous models that have been used to conceptualize visual STM and WM ability. For example, some researchers take a “slots-based approach”. In this conceptualization, capacity is limited in terms of slots that store a singular memory representation for each individual item (e.g. Cowan, 2005; Luck & Vogel, 1997; Pashler, 1988). A memory representation is either in memory, or it is absent from memory. This model has been challenged by studies involving the precision of the contents of memory. Researchers utilizing this “resource-based approach” view the contents of memory in terms of resolution. For example, several researchers have recently provided evidence that the quality of the memory representation depends on the number of items that are being held in STM or WM (Alvarez & Cavanagh, 2010; Awh, Barton, & Vogel, 2014; Ma, Husain, & Bays, 2014).

Yet another important consideration when measuring visual STM or WM is the type of task that researchers are using to generate these capacity estimates. Recently, for example, change-detection tasks have come under a bit of scrutiny as several researchers have suggested that performance does not strictly reflect the storage of individual items and that it is possible that there is some sort of ensemble or configural processing that is occurring (Brady & Tenenbaum, 2013;

Pailian & Halberda, 2015; Vogel & McCollough, 2005; or Brady, Konkle, Alvarez, 2011 for a review). For example, it’s possible that adults may be able to detect an item has changed color in some general way, but cannot describe *what* that change was, or *where* it occurred. This would not be possible under strict “slot” theories of capacity (i.e. how many individual items are in memory), as ostensibly each encoded item should result in 100% recall. It is also important to note that while configural processing is a distinct possibility, meaningful individual difference measures are still likely to occur and perhaps a more “continuous” approach (i.e. a measure of the strength or quality of the memory representation) to visual STM quantification may allow us to assess the strength of individual memory representations independent of button-presses. This could eliminate noise driven by errors due to poor response mapping or perseverative button presses, and enable a more continuous measure of memory, free from assumptions regarding slots vs. resource models of STM.

Physiological Measures of STM: Task-Evoked Pupillary Responses

In addition to these behavioral measures of capacity, physiological measures hold promise for assessing STM capacity. For example, changes in task evoked pupillary responses (TEPRs) have long been used as an index of perceptual and cognitive processing (Beatty, 1982; Heaver & Hutton, 2011; Hess

& Polt, 1964; Unsworth & Robison, 2015). Though pupils change in response to ambient light (Goldwater, 1972), researchers have now firmly established that pupils also show task related changes for numerous memory processes including visual STM maintenance (Unsworth & Robison, 2015), recognition memory (Goldinger & Papesh, 2012; Heaver & Hutton, 2011), and novelty detection (Sirois & Jackson, 2012). For example, recent work with adults demonstrates that pupil response during the retention interval of a change-detection task was significantly modulated by the number of items that were held in the participants visual STM (Unsworth & Robison, 2015). This suggests that changes in pupil size may be a marker of STM maintenance. Moreover, for set sizes *within visual STM capacity*, pupil size initially decreased and then increased over the retention interval. In contrast, when participants were asked to remember a number of items *at or above their capacity*, pupil dilation initially increased, then remained high for the remainder of the retention interval. Importantly, this relationship varied as a function of each individuals' maximum visual STM capacity score (Unsworth & Robison, 2015). These findings demonstrate the utility of using pupil size as an index of STM maintenance.

Though Unsworth & Robinson (2015) did not report pupil changes during the *test* array, others have demonstrated changes in pupil size for both visual recognition memory, and novelty detection with adult participants (Goldinger & Papesh, 2012; Heaver & Hutton, 2011) and for familiarity (Gredebäck, Eriksson, Schmitow, Laeng, & Stenberg, 2012; Hellmer, Soderlund, & Gredeback, 2016) and novelty detection (Sirois & Jackson, 2012) in infant participants. For example, using a visual recognition task with adult participants, Heaver and Hutton (2011) found pupils to be significantly larger in response to familiar stimuli than to novel ones. Similarly, Hellmer and colleagues (2016) found that both 7-month-old infants and adults showed significantly larger pupil dilation to familiar items and significant constriction to novel items — after a brief familiarization period. Hence, this literature supports the idea that pupils dilate to familiar items and constrict to novel ones. Note that although auditory "oddball" paradigms typically produce dilation to novel and deviant tones (Wetzel, Buttelmann, Schieler, & Widmann, 2016), this is likely driven primarily by arousal responses rather than cognitive responses, *per se*.

Infant Visual Short-Term Memory: Techniques and Measurement

Previous work with infant participants demonstrates that like adults, infant visual STM appears to be highly capacity limited (e.g., Ross-Sheehy, Oakes, & Luck, 2003; Oakes, Hurley, Ross-Sheehy, & Luck, 2011; Oakes et al., 2009; Rose, Feldman, & Jankowski, 2001). These limitations constrain the amount of visual information that an infant can keep active in memory and subsequently limits the kinds of things that an infant can learn. For example, being able to hold the properties of one toy in mind while examining a second toy should facilitate the detection of both shared features and unique features, which are important categorization and learning cues (Oakes et al., 2011). Additionally, capacity limits are significant because they may force the infant to adopt other approaches to dealing with the enormous amount of visual information that they encounter on a daily basis. For example, research conducted by Feigenson (2007) suggests that children as young as 14 months of age are capable of chunking items together when they are presented in groups, presumably lessening the memory load.

Measuring visual STM in infants, however, has proven difficult, especially doing so in a way that isolates visual STM in infancy the same way this system has been isolated in adults. For example, early infant STM memory studies typically relied on some form of habituation or familiarization (Blaga & Colombo, 2006; Fagan, 1984; Rose, Feldman, & Jankowski, 2002). However, these familiarization tasks failed to isolate STM from longer-term memory (LTM) mechanisms, making specific claims about the development of STM problematic. According to Baddeley and colleagues, STMs are formed very rapidly (e.g., milliseconds), are highly capacity limited, and decay rapidly in the absence of explicit rehearsal (Baddeley & Logie, 1999; Baddeley & Hitch, 1994). Thus, tasks that rely on familiarization are likely tapping longer-term memory systems.

In an attempt to isolate visual STM, Ross-Sheehy, Oakes, and Luck developed a modified change-detection task, based on the paired-comparison technique (Oakes et al., 2006; Oakes et al., 2009; Ross-Sheehy et al., 2003; Ross-Sheehy, Oakes, & Luck, 2011). Specifically, Ross-Sheehy et al. (2003), presented infants with two separate arrays (on two separate monitors) of 1, 2, or 3 different colored squares. These arrays appeared on each screen for 500ms, then disappeared for 250ms, then reappeared for 500ms, and so on (see Figure 2). This cycle repeated for 20s. On one monitor, a single randomly selected square changed color each time that the display reappeared (i.e., "change" array). On the other monitor, the colors remained the same for each reappearance (i.e., the "no-change" array). Ross-Sheehy et al. (2003) reasoned that if an infant can rapidly encode the items on the changing array and retain that memory across the 250ms retention interval, they should prefer the changing array and look systematically longer to that side. To assess change detection, change preference was calculated as the proportion of looking to the change array relative to looking for both the change and no change arrays. These proportion scores were then compared to chance (.50), and any significant deviation above chance was taken as evidence of change detection. Results revealed significant change preferences for set size 1, 2, and 3 for the 10- and 13-month-old infants, whereas 4- and 6-month-old infants demonstrated significant change preference only at set size 1. This work was the first to demonstrate that 1) infants can rapidly encode items into visual STM; 2) this ability is apparent as early as 4 months; and 3) there is a significant increase in visual STM for object identity (color) somewhere between 6 and 10 months of age. Important follow-up work narrowed this window even more, suggesting the bulk of this development

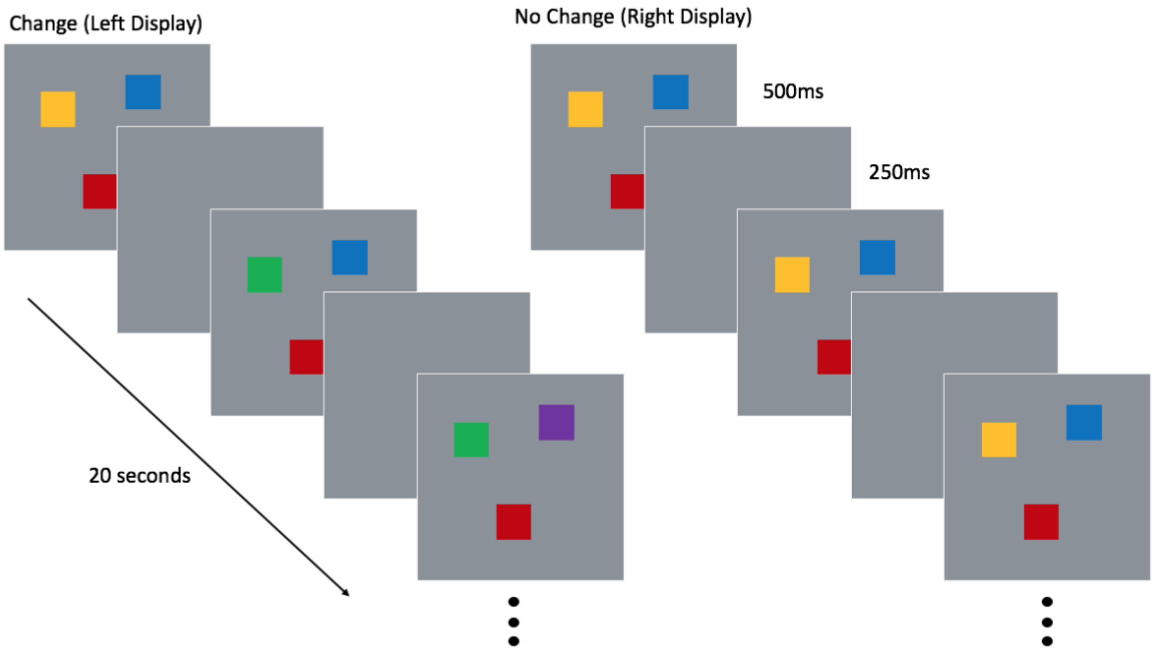


Figure 2: Schematic Representation of a Trial Adapted from Ross-Sheehy et al., 2003

occurred between the ages of 8 and 9 months (Oakes, Ross-Sheehy, & Luck, 2006; Oakes, Messenger, Ross-Sheehy, & Luck, 2009). This basic pattern of development was found for other forms of memory, including memory for object location (Oakes, Ross-Sheehy, & Luck, 2007; Oakes, Hurley, Ross-Sheehy, & Luck, 2011)and non-linguistic auditory information (Ross-Sheehy & Newman, 2015).

What might be driving these rapid improvements in memory? What else is undergoing rapid change during this narrow window? One possible explanation is that improvements in attentional shifting that accompany parietal lobe development may account for these performance improvements (Johnson, 2008; Posner & Dehaene, 1994). This makes sense, as previous control studies demonstrated that infants younger than 8 months of age fail to detect color changes, even when every single square in the array changes color (Oakes et al., 2006; Oakes et al., 2009). This is striking, as infants need to only encode a single item to detect the change. However, all infants younger than 8 months fail at this task. This suggests that multi-item arrays are very attentionally demanding, and failures of change detection may be driven by failures of selective attention. To assess this more directly, Ross-Sheehy, Oakes, and Luck (2011) showed 5- and 10-month-old infants arrays of squares that exceeded capacity (i.e., set size 6 arrays). Arrays blinked on and off (500ms on, 300ms off, and so on), for a total of 20 seconds. However, unlike previous work, all of the arrays were “change” arrays and were shown on a single screen. What varied, was the location of a spatial attentional precue—it either appeared in the location of a color change (valid) or in the location of one of the non-changing squares (invalid). Results demonstrated that by adding a spatial attention cue, infants were able to detect the color change, in arrays that were otherwise beyond capacity (Ross-Sheehy, Oakes, and Luck, 2011).

Though paradigms like these have demonstrated some visual STM ability in infants as young as 4 months, there have been some important critiques of these procedures. For example, some have speculated that the dual-stream procedure may be tapping multiple memory mechanisms, and thus, infant performance may not be directly comparable to performance on change-detection tasks used in toddlers and adults (Simmering, 2011; Simmering & Perone, 2013). Indeed, the authors clearly state that infants may significantly prefer changing streams even when they contain set sizes above that infant's capacity (a possibility that is also noted in Ross-Sheehy, Oakes and Luck, 2003). This may be possible, for example, as infants have 20s to build up memories for the non-changing items. Thus, it may be the case that infants grow bored of the non-changing array, resulting in a preference for the changing array. Because there is no explicit behavioral response in these tasks, it is not at all clear what is driving the behavior. Moreover, because infant variants of the change-detection tasks are typically *qualitatively different* than adult change-detection tasks, it is currently unclear if these early works are assessing similar mechanisms tested using adult change-detection tasks.

Recent work aimed at addressing this very problem (Oakes, Baumgartner, Barrett, Messenger, & Luck, 2013) has utilized eye tracking in an attempt to measure change detection within a given array (more similar to adult approaches). In this task, 6- and 8-month-old infants viewed a sample array, containing two single feature items, for 517ms. This was followed by a 317ms retention interval and then a 3000ms test array, consisting a familiar color and a novel color (see Figure 3). Results suggest that the 8-month-old infants demonstrated a significant preference for the changed item, suggesting that they have some memory for colors; however, the 6-month-old infants provided no such evidence. In another variation of this experiment, both items in the sample array changed (i.e. a red and green item during the sample array was replaced by a blue and yellow item during the test array). In this instance, both the 6- and 8-month-old infants showed a change preference. This work constitutes an important intermediate step between the dual stream procedures typically utilized in infants, and single screen procedures typically used in adults. Further, these results are the first to demonstrate that infants are capable of encoding information into their visual STM after a single brief exposure (Oakes et al., 2013).

In an attempt to assess visual STM across a larger variety of set sizes, Ross-Sheehy and Eschman (2019) recently developed a modified change-detection task modeled after adult change-detection procedures (Luck & Vogel, 1997). In this task, infant and adult participants were shown arrays of 1-4 colored

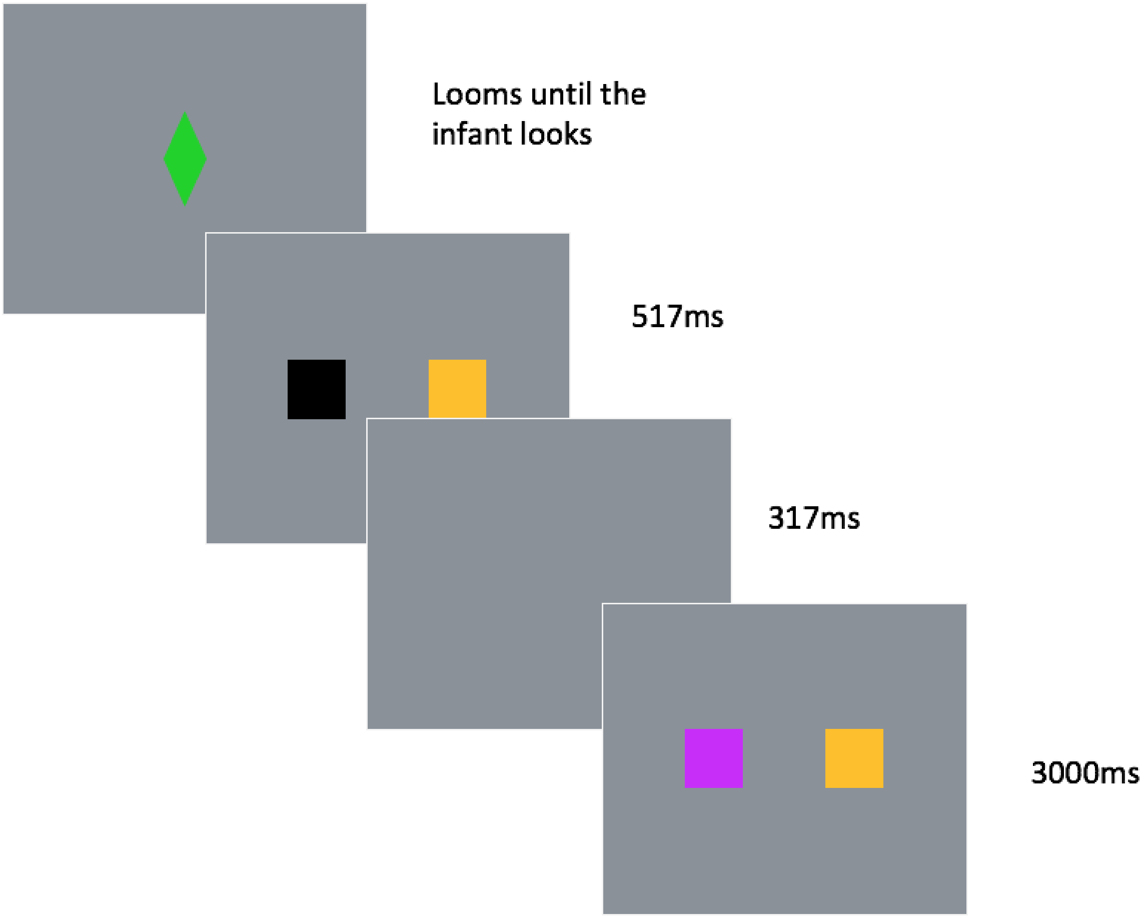


Figure 3: Schematic Representation of a Trial Adopted from Oakes et al., 2013

circles presented using a “one-shot” protocol (i.e., one sample array followed by one test array). The test array was either exactly the same (no-change trial) or one of the circles randomly changed color (change trial, see Figure 4). Both infants and adults participated in the same task with two minor exceptions. First, in addition to collecting eye tracking measures, adults indicated "same" or "different" via button press, and those responses were used to calculate accuracy. Second, while both infants and adults were given feedback following the “change” test arrays, infants saw a “reward” animation in the location of the color change, whereas adults heard a tone if they responded incorrectly. Change preference scores in both infant and adult participants were calculated as the proportion of time spent looking to the changed circle relative to looking to all the circles during the test array. This is a unique feature of the “one-shot” methodology, as previous infant change detection tasks (Oakes et al., 2013; Ross-sheehy et al., 2003), calculated change preference based on looking across two distinct displays. Results revealed interesting developmental effects. Specifically, 5- and 8-month-old infants had significant change preferences for set-size 2 only, 11-month-old infants had a significant change preference for set-size 2 and 3, and adults had a significant change preference for set-size 2, 3, and 4 (Figure 5). However, these results should be interpreted with caution as a follow-up analysis examining the effect of the previous fixation revealed that change preference was almost entirely driven by chance looking to the “to-be-

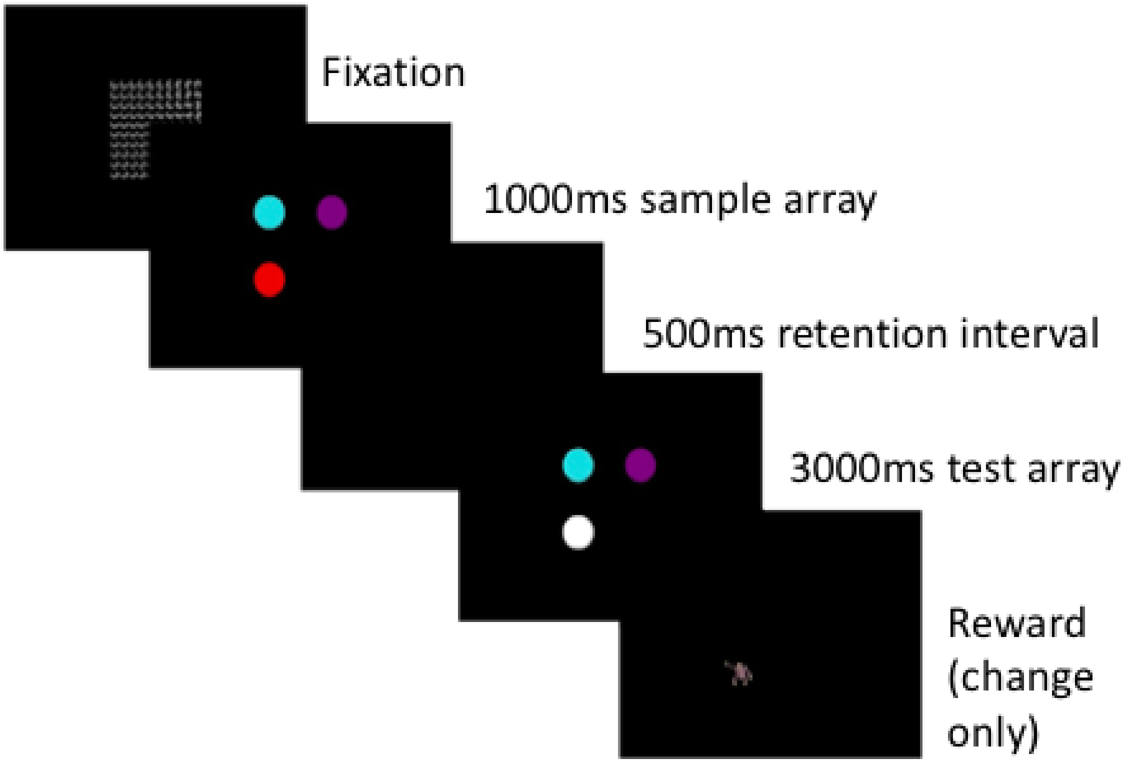


Figure 4: Schematic Representation of a Trial Used in Ross-Sheehy & Eschman, 2019

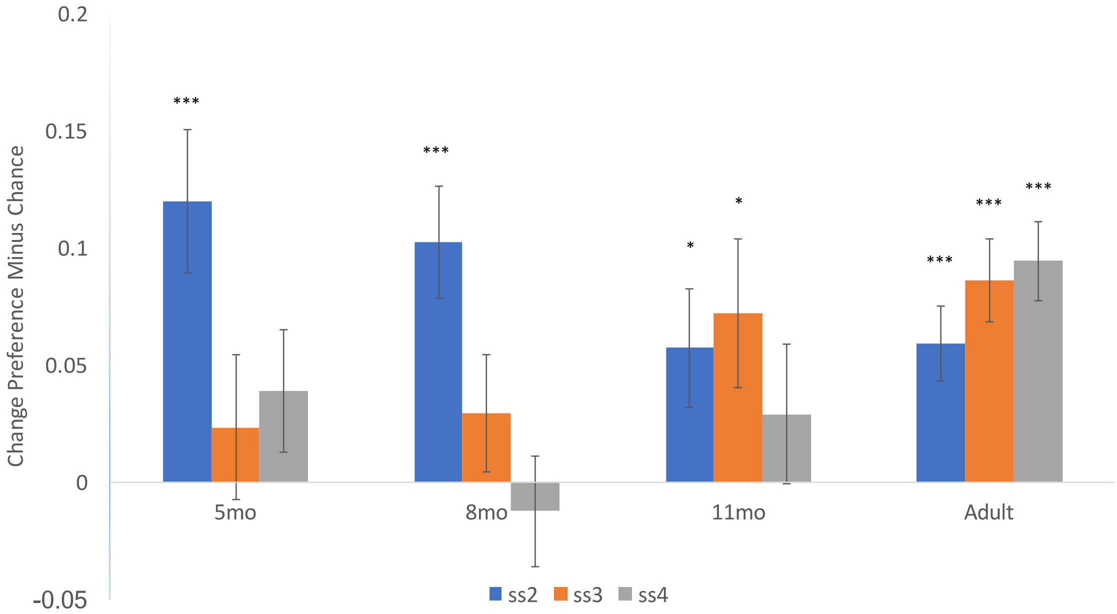


Figure 5: Change Preference Minus Chance for Each Set Size, as a Function of Age and Set Size (Ross-Sheehy & Eschman, 2019)

changed” item during the encoding array (Figure 6). Specifically, if infants and adults happened to fixate the “to-be-changed” item during the encoding array, they had a significantly higher change preference than when they did not. Thus, they concluded that change preference in the context of a one-shot, change-detection task, may not be an accurate measure visual STM capacity. However, these results do clearly demonstrate capacity effects in terms of recency – items viewed during the encoding array were retained across the retention interval and influenced subsequent preference for the changing item during the test array. If we can manipulate the temporal dynamics of these encoding fixations, it may be possible to estimate capacity based on change preferences as a function of recency effects. This will be tested in the current project.

In addition to these change preference scores, Ross-Sheehy and Eschman (2019) also report some interesting physiological results that support previous research. Specifically, pupils were significantly larger for *familiar* (no-change) trials than for *novel* (change) trials. These findings are consistent with previous work (Heaver & Hutton, 2011) and suggests that pupil dilation in this task may be driven by recognition responses, rather than *change*-detection (Ross-Sheehy & Eschman, 2019). In addition, task-evoked pupil responses (TEPRs; i.e., pupil change from baseline) were the largest for set size 1 and decreased in size as set size increased over the course of the test array, suggesting some cognitive processing that varied as a function of set size.

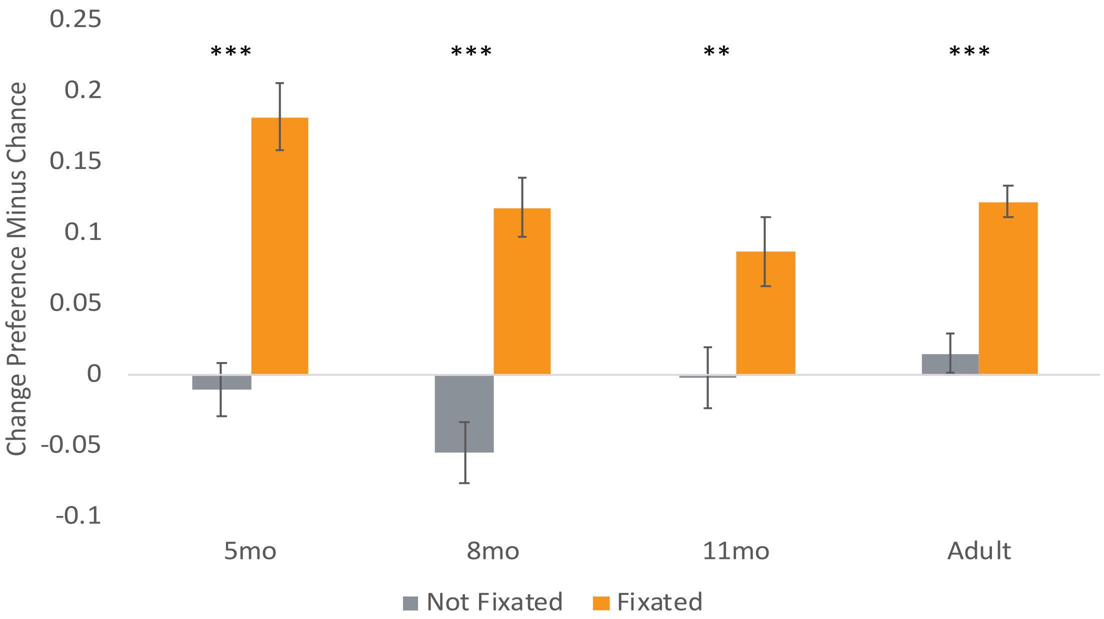


Figure 6. Change Preference as a Function of Prior Fixations to the To-Be-Changed Item During the Sample Array

Current Project

The goal of the current project was to measure visual STM in infants, using a similar approach used in adult STM tasks, in hopes of finding a more nuanced way of quantifying individual differences in visual STM ability. Additionally, this project aims to clarify the relationship between infant measures of visual STM and adult visual STM. To accomplish this, this project incorporated several novel change-detection tasks that replicate and extend previous infant and adult visual STM work. Specifically, Experiment 1 identified visual dynamics (e.g. low-level visual behaviors and pupil changes) in adults that were related to accuracy in a classic change-detection task. Experiment 2 utilized these behavioral and physiological correlates of visual STM and found evidence of memory maintenance in 5- and 11-month-old infants. Experiment 3 measured the extent to which modifications in testing procedures for infant participants (e.g., larger eccentricities, longer trial durations, etc.) generally influenced measures of visual STM capacity in adults. Finally, Experiment 4 determined the extent to which change preference varies as a function of previous attention/fixation. Combined, these four experiments have provided a new means to assess visual STM in adults in addition to helping identify key markers of visual STM in infants. Specifically, they have helped determine if things like duration and size influence memory traces. These experiments have also helped to identify visual dynamics that correspond with successful memory maintenance and helped to clarify the role of fixation order in change detection. Importantly, this project has contributed significantly to the overall understanding of the relationship between infant visual STM and adult visual STM/WM measures.

CHAPTER TWO

EXPERIMENT 1: ASSESSING THE BEHAVIORAL AND PHYSIOLOGICAL CORRELATES OF STM MAINTENANCE IN ADULTS

The goal of Experiment 1 was to identify pupil and eye movement correlates of visual STM performance (response accuracy) in adults. To accomplish this, a classic adult change-detection task (Luck & Vogel, 1997), was replicated with the added precision of eye tracking (Figure 7). This allowed for the collection of gaze and pupil behaviors while adults were performing the task. As mentioned above, Ross-Sheehy and Eschman (2019) demonstrated a meaningful connection between visual behavior, pupil changes, and accuracy. The current task, building off of these results, incorporated set sizes that were well-beyond typical adult capacity (i.e. > 4 items). Thus, it is possible to identify eye movement patterns and pupil responses both for correct and incorrect responses. In the Luck and Vogel (1997) task, participants briefly viewed an array of 1, 2, 3, 4, 8, or 12 colored squares (100ms), followed by a retention interval (900ms), and then a second *test* array (2000ms). The test array was either identical to the first (i.e., "no change"), or one randomly chosen square changed color ("change"). Previous findings demonstrate that adults could remember the features of 3-4 objects (Luck & Vogel, 1997).

Based on previous work, it was expected that adult performance would decline as set size increased, and visual dynamics (i.e., saccades, fixations, and

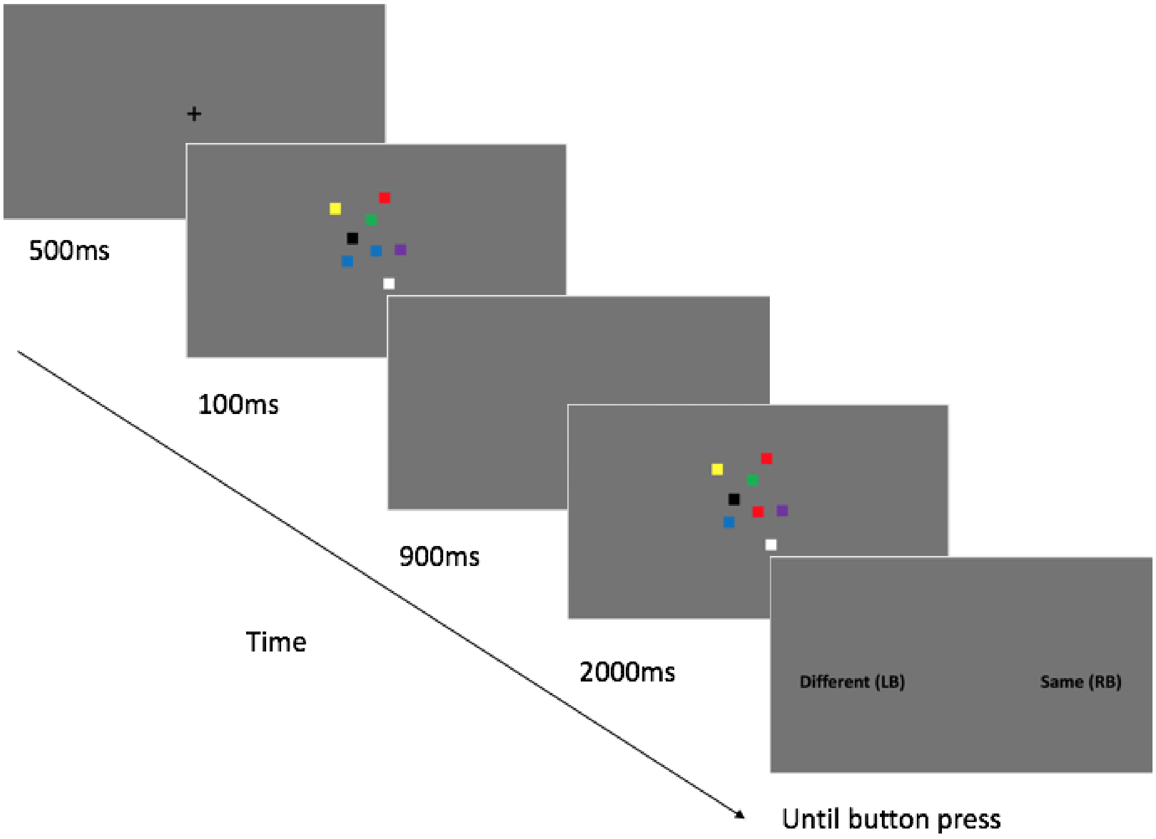


Figure 7: Experiment 1, Change Detection Task (Luck & Vogel, 1997), Set Size 8 Change Trial

pupil size) would be related to response accuracy. Specifically, by examining saccades, fixations, and pupil changes in response to “correct” and “incorrect” trials, it is possible to identify moment-to-moment changes in visual behavior that were related to behavioral accuracy. The relationship between trial-to-trial variations in visual STM accuracy and visual dynamics has never been examined and may potentially inform theories regarding individual differences in visual STM maintenance and capacity in adults. Additionally, this method may help us identify similar patterns of responding in populations where overt behavioral responses are impossible (e.g. infants).

Method

Participants

A total of 25 participants completed a replication of the Luck and Vogel (1997) change-detection task. One participant’s data was removed from the final data set due to difficulty during calibration that resulted in poor eye tracking data quality (lost gaze on more than 50% of the trials). The final sample consisted of 24 adults (17 females, 6 males, and one participant chose not to select a gender, *M* age = 20.88 years). Ethnicities were reported as follows: White (n=18), Black (n=2), and Asian (n=4). All participants had normal or corrected to normal acuity and were screened for colorblindness (Ishihara, 1960). Additionally, all participants reported no serious head injuries or neurological disorders. Adult participants were recruited from the University of Tennessee’s online recruitment tool and were awarded course credit for their participation.

Apparatus

Adult Eye Tracking Setup: An Eyelink 1000+ eye tracking system with a remote desktop mount and 25mm lens was used to collect data for the adult participants. An 890nm infrared light emitter was used to measure continuous gaze. Point of gaze data and pupil area were sampled monocularly at 1000hz, and adults were calibrated using a 5-point calibration scheme. All eye tracking data were gathered from the participants left eye. Pupil size was measured in terms of total area. Adults were tested in a dimly lit room and sat approximately 65cm in front of high-performance 24”, 120Hz, Asus monitor (native resolution of 1920x1080), with a viewable surface of 45.5° (w) by 26.76° (h). Participants used an Xbox gaming controller to indicate a “same” (right bumper) or “different” (left bumper) response.

Stimuli

As in Luck and Vogel (1997), trials consisted of arrays of 1, 2, 3, 4, 8, or 12 colored squares randomly selected from a series of 7 highly discriminable colors (red, purple, blue, white, black, green, and yellow; RGB and luminance values can be found in Table 10 *in the appendix*). Colors were drawn randomly with replacement but with the contingency that no color appears more than twice in a given array. Each square measured 0.65° and arrays were constrained to appear within a 9.8° X 7.3° region with at least 2° of separation between each square (center to center). Each trial started with a 100ms *sample array* and was followed immediately by a 900ms blank *retention interval*, and finally a 2000ms *test array*. The 900ms blank retention interval is important as it ensures that that the contents of memory are beyond sensory memory and have transitioned in the STM (Luck & Vogel, 1997; Phillips, 1974). For half of the trials, the test array was exactly the same as the sample array (i.e. no change trials), and for the other half, the color of a single, randomly chosen square changed (i.e. change trials).

Design and Procedure

This task incorporated a 2 x 6 design with both condition (change and no change) and set size (1, 2, 3, 4, 8, and 12) manipulated within subject. Participants were seated approximately 65cm in front of a high-performance 24” 120Hz monitor (native resolution of 1920x1080), with a viewable surface of 45.5° (w) by 26.76° (h). Participants saw up to 40 blocks of 12 randomized trials, and each block contained one of every possible trial type for a max of 480 trials. Not all participants were able to complete all 480 trials in the allotted 60m block. The experimenter was seated behind a black curtain and monitored eye movements from a small video monitor.

Participants were instructed to fixate a small black fixation cross located in the middle of the screen to initiate the trial presentation. Once they fixated this cross for 500ms, the trial would begin. As in Luck & Vogel (1997) each trial consisted of a sample array, followed by a retention interval, and finally a test array (see *Stimuli* for details). Following the test array, participants were instructed to select “same” by hitting the right bumper on the gaming controller if they believed the sample and test arrays were the same, and “different” by hitting the left bumper if they believed sample and test arrays were different.

Primary behavioral dependent measures included accuracy and measures of capacity (Pashler, 1988). In addition, several eye tracking measures were calculated, including the number of saccades, run count, average fixation durations, and TEPRs, as previous work (Ross-Sheehy & Eschman, 2019) suggested that these measures were likely to be related to behavioral measures of accuracy and capacity in adults. Trials where adults failed to provide a behavioral response (i.e. did not hit the button in time) were removed from data analysis. Additionally, due to the near ceiling effects for set sizes < 4, the following series of analyses (other than percent correct) were conducted only on set sizes 4, 8, and 12. Additionally, the effects presented below are all visual metrics that were gathered during the “test” period of each trial. It is also important to note that interest areas (IAs) were place around each individual item in both the sample and test array. The IAs were slightly larger than the items (+10 pixels in height and +10 pixels in width).

Results

Eye Tracker Event Parsing

For this and all subsequent analyses, eye tracker event parsing was as follows: Saccades, fixations and blinks were calculated online using EyeLink standard online event parser (Stampe, 1993), which incorporates a velocity threshold algorithm to classify saccades with the following settings: Saccade velocity > 40°/s and acceleration > 8000°/s for a minimum duration of 8ms. Samples that do not exceed these thresholds are classified as either fixations (pupil data present) or blinks (no pupil present for at least 3 consecutive samples). Nearby fixations were left unmerged.

Percent Correct

Mean percent correct was calculated for each condition at each set size. Performance was near ceiling for set sizes 1-4, then dropped precipitously as set size increased (Figure 8). To assess accuracy as a function of condition and set size, a repeated measures ANOVA was conducted with change status (change, no change) and set size (1, 2, 3, 4, 8, & 12) as within subjects variables. Results revealed a significant main effect of set size $F(5,115)=131.718$, $p<.001$, $\eta^2_p=.851$, suggesting that task difficulty increased as set size increased. Additionally, there was a significant main effect of change status $F(1,23)=32.792$, $p<.001$, $\eta^2_p=.588$, suggesting that in general, performance was higher for “no change” trials. Both of these effects were subsumed under a significant Set Size x Change Status interaction $F(5,115)=35.150$, $p<.001$, $\eta^2_p=.604$. As set size increased, so too did the difference between the “change” and “no change” trials. Follow-up simple effects tests with a Sidak correction for multiple comparisons showed a robust significant difference in percent correct as a function of change status at the largest set sizes, set size 8 ($p<.001$), and set size 12 ($p<.001$) as participants are more accurate on no change trials. This is to be expected as participants are more likely to assume that something did not change if they fail to notice or are unsure if a change that has taken place. Additionally, due the significant lack of incorrect responses for set sizes 1-3 (Figure 8), only set size 4, 8, and 12 were included in the following analyses.

Saccade Count

Saccade count was calculated as the average number of eye movements or *saccades* that an individual made during each test array. Mean scores were then analyzed using a repeated measures ANOVA with change status (change, no-change), set size (4, 8, 12) and response accuracy (correct, incorrect) as within subjects variables. Results are summarized in Table 1. Not surprisingly, results revealed a significant main effect of set size $F(2,22)=6.751$, $p=.005$, $\eta^2_p=.380$, suggesting that individuals made more saccades on trials with larger set sizes. Follow-up simple effects tests with a Sidak correction for multiple comparisons show that there is a significant difference in saccade count between set size 4 and 12 ($p= .029$) and between set size 8 and 12 ($p=.012$). Importantly,

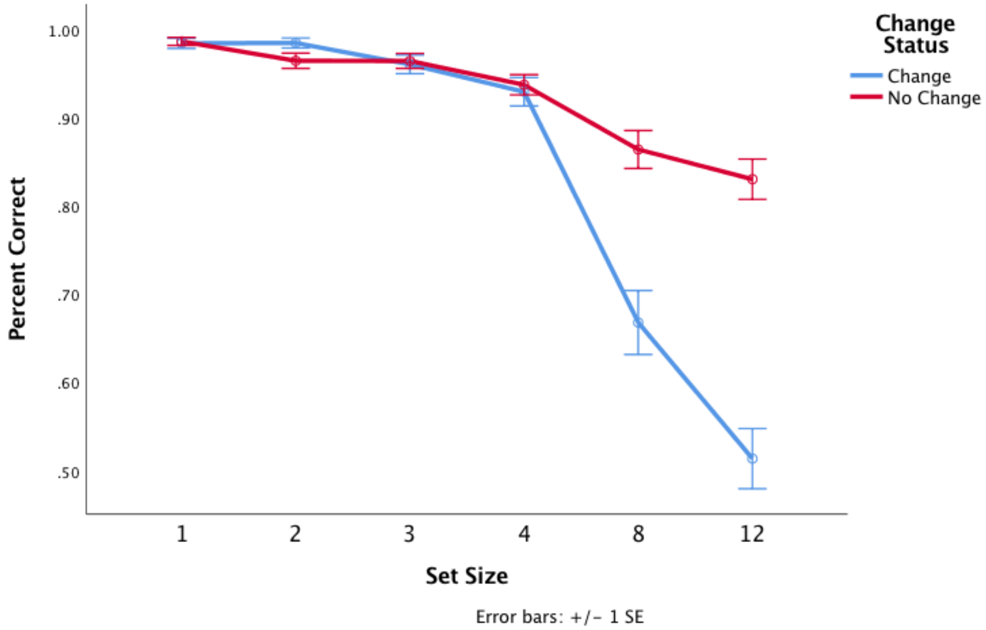


Figure 8: Average Percent Correct as a Function of Change Status and Set Size.

Table 1: Analysis of Variance Visual Dynamics Results (Experiment 1)

	F	*p*	*df*	df_{error}	η^2
SACCADE COUNT					
Set Size	6.751	***0.005***	2	22	0.38
Change Status	0.716	0.416	1	11	0.061
Accuracy	8.976	***0.012***	1	11	0.449
Set Size X Change Status	1.057	0.364	2	22	0.088
Set Size X Accuracy	0.308	0.738	2	22	0.027
Change Status X Accuracy	1.792	0.208	1	11	0.14
Set Size X Change Status X Accuracy	0.395	0.679	2	22	0.035
RUN COUNT					
Set Size	40.198	***<.001***	2	22	0.785
Change Status	0.074	0.791	1	11	0.007
Accuracy	0.261	0.619	1	11	0.023
Set Size X Change Status	0.893	0.424	2	22	0.075
Set Size X Accuracy	0.517	0.604	2	22	0.045
Change Status X Accuracy	0.04	0.846	1	11	0.004
Set Size X Change Status X Accuracy	1.009	0.381	2	22	0.084
AVERAGE FIXATION DURATION					
Set Size	1.474	0.198	2	22	0.137
Change Status	0.101	0.756	1	11	0.009
Accuracy	7.827	***0.017***	1	11	0.416
Set Size X Change Status	1.132	0.34	2	22	0.093
Set Size X Accuracy	1.406	0.266	2	22	0.113
Change Status X Accuracy	0.078	0.785	1	11	0.007
Set Size X Change Status X Accuracy	1.797	0.19	2	22	0.14
TEPRs					
Time (bins)	18.974	***<.001***	3	72	0.442
Accuracy	5.994	***0.022***	1	24	0.2
Bin X Accuracy	0.831	0.481	3	72	0.033

results also revealed a significant main effect of accuracy, *F*(1,12)=8.976, *p*=.012, η^2_p=.449, revealing that participants made significantly more saccades when viewing test arrays prior to an incorrect response (Figure 9). These results suggest that saccade count may be an informative metric when it comes to measuring visual STM performance.

Run Count

Run count, which is related, but different from saccade count, is defined as the number of looks to individual squares. Note that consecutive looks to the *same* square does not increment the run count, although looking back and forth between even just two squares does. The measure allows researchers to gauge the degree to which the participant scanned unique array items, presumably reflecting between-item comparisons during the test interval. This measure provides additional insight into the process(es) involved during the test array. Specifically, run count adds a precise measure of how many times participants looked at each individual item. This is important as differences in run count as a function of change status, set size, and response accuracy may be speaking to the underlying process(es) utilized by participants while preforming the task. Mean run count was analyzed using a repeated measures ANOVA with change status (change, no-change), set size (4, 8, 12), and response accuracy (correct, incorrect) as within subjects variables. Results revealed a significant main effect of set size, *F*(2,22)=40.198, *p*<.001, η^2_p=.785, suggesting that individuals

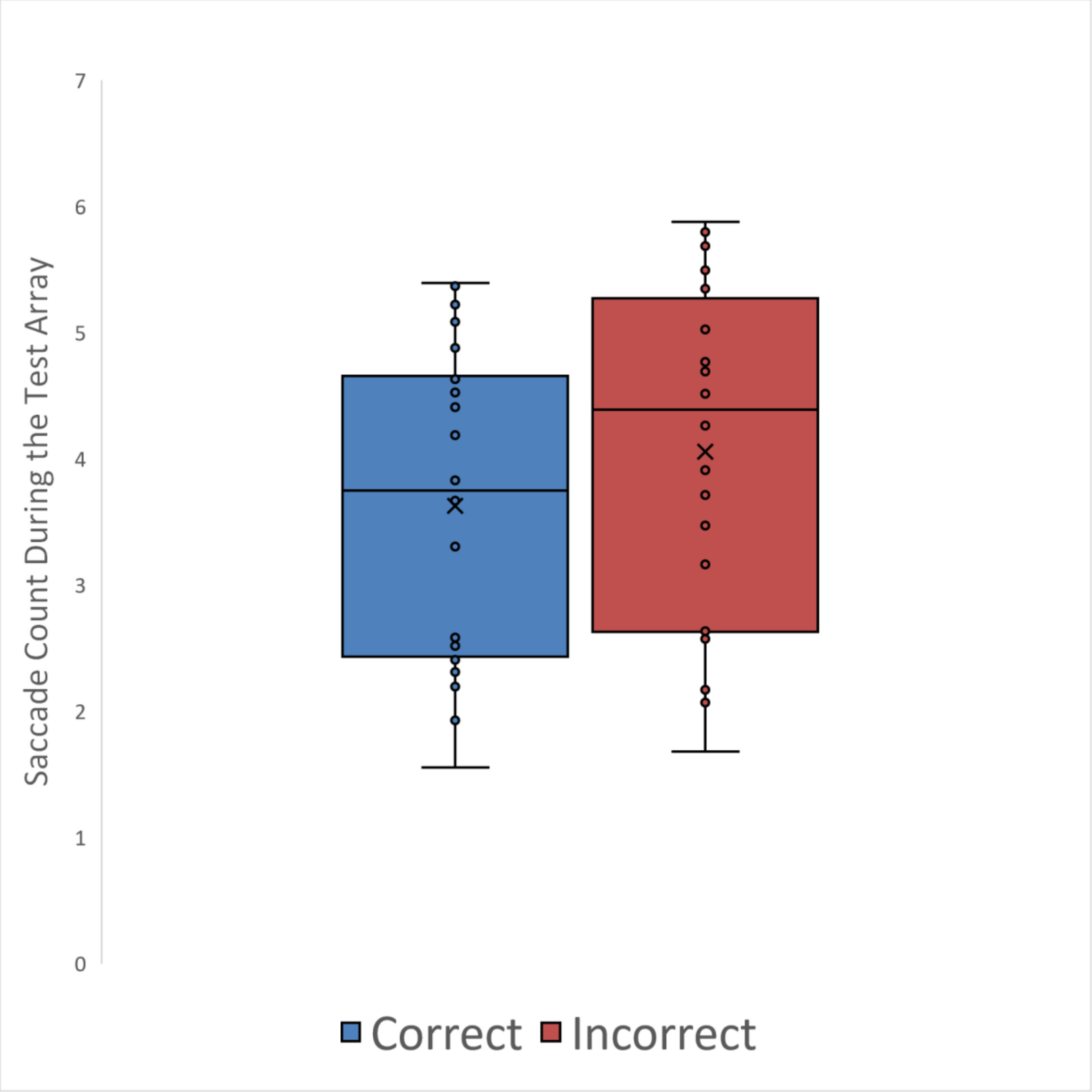


Figure 9: Saccade Count as a Function of Response Accuracy and Change Status During the Test Array

engaged in increased scanning on trials when set sizes were larger. Follow-up simple effects tests with a Sidak correction for multiple comparisons show that there is a significant difference in run count for set size 4 and 8 (p=.005) as well as for 8 and 12 (p<.001), and for 4 and 12 (p<.001), again demonstrating that participants had longer run counts for larger set sizes. Although run count did not vary as a function of change status or accuracy, this is most likely due to the fact that the total eccentricity was very small, approximately 9.8° X 7.3° .Thus, to perform well, adults did not need to scan multiple items, as all were reasonably close to central fixation. Experiment 3 will explore this hypothesis, by substantially increasing the total eccentricity of arrays.

Average Fixation Duration

Average Fixation Duration was calculated by computing the average duration (in milliseconds) of each individual fixation during the test array. Again, means were analyzed via a repeated measures ANOVA with change status (change, no-change), set size (4, 8, 12) and response accuracy (correct, incorrect) as within subjects variables. While there was no main effect of set size or change status, results revealed a main effect of response accuracy, $F(1,11)=7.827$, $p=.017$, $\eta^2_p=.416$ (Figure 10). This is important and suggests that longer fixations are either directly contributing to higher accuracy, or that individuals who tend to fixate longer, tend to have higher WM ability. Taken

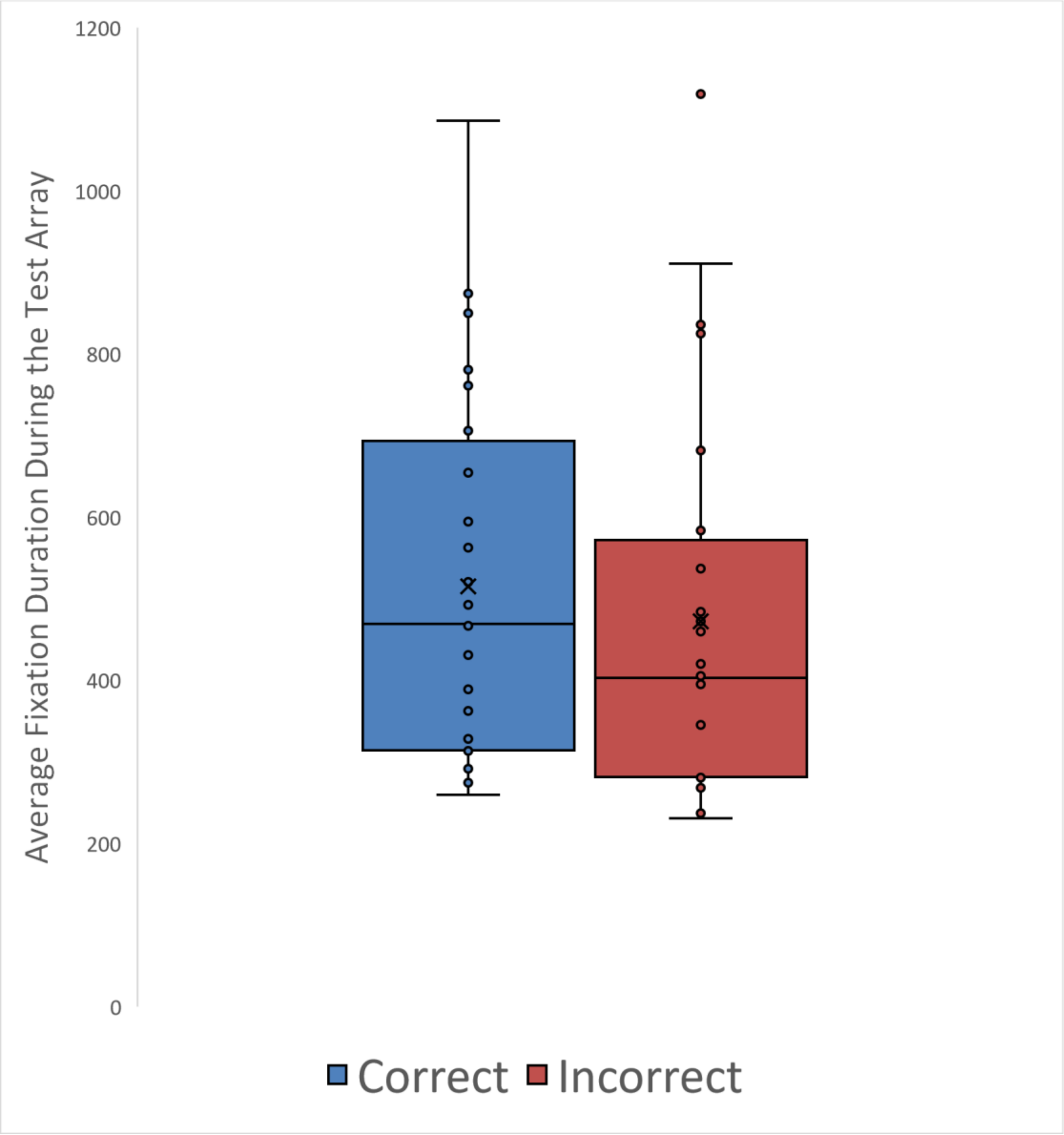


Figure 10: Average Fixation Duration as a Function of Response Accuracy

together with saccade counts, these results demonstrate that in general, adults made more eye movements and had shorter average fixation for incorrect trials compared to correct trials.

Capacity Estimates

In an attempt to connect this research with previous conceptualizations of visual STM, capacity (K) for each group was estimated using Pashler's (1988) formula. The formula for Pashler's K was as follows: $K = SS * (H-FA) / (1-FA)$ based on the hit (H) and the false alarm (FA) rates for each set size (SS), with H calculated as: Hits/(Hits + Misses) and FA calculated as: False Alarms/(False Alarms + Correct rejections). Note that this formula assumes all FA are guesses, and so corrects the estimate accordingly. Although this correction may not be appropriate for all participant populations (e.g., toddlers are more likely to perseverate on "different" responses), it is used here to make contact with existing literatures and provide a means of assessing individual differences. This formula is typically utilized for estimating capacity for the whole report version of the change detection task, as was used here. Pashler's K estimate was 4.67 (*SD* = 1.79) for set size 8 and 4.79 (*SD* = 2.43) for set size 12. These scores were averaged together to yield a single K estimate (4.73). To assess the relation of K estimates to the visual dynamics identified above, a bivariate correlation revealed that Pashler's K (4.73) was significantly negatively correlated with saccade count and run count and positively correlated with average fixation duration (see Table 2).

Table 2: Capacity Estimate and Visual Dynamic Correlations (Experiment 1)

Measures	1.	2.	3.	4.	5.
1. Pashler's K	1	.771**	-.046**	-.159**	.162**
2. d-Prime		1	-.103**	-.146** .	171**
3. Run Count			1	.410**	-.238**
4. Saccade Count				1	-.768**
5. Average Fix. Duration					1

** Correlation is significant at the .01 level

In addition to Pashler’s K estimates described above, d-prime (d’) was calculated for each participant, to help provide an assessment of memory strength that was agnostic with respect to theoretical differences in models of working memory (Wickens, 2001). In the context of the present task, greater sensitivity (i.e., higher d’ scores) suggest greater change detection. Similar to the previous estimate of capacity, d’ was calculated for set size 8 was 1.782 (*SD* = .97), and for set size 12 was 1.319 (*SD* = 1.15). These scores were averaged together to yield a single d’ estimate (1.55). Again, a bivariate correlation revealed that d’ was significantly negatively correlated with saccade count and run count, and positively correlated with average fixation duration (Table 2). Although the magnitude of these effects is small, this relation suggests that these visual dynamics are capable of identifying meaningful individual differences in visual STM performance. Further, they suggest that low-level visual dynamics such as run count, saccade count, and fixation duration vary with behavioral accuracy. Due to the nature of this task (100ms encoding period), it is currently not possible to know if these correlations represent trial to trial fluctuations in performance or effort, or something more stable, like general STM ability. This question will be explored in more detail in Experiment 3.

Pupil Analyses

Raw pupil sizes during the 2000ms test array were baseline corrected using the first 100ms of each sample array. This eliminated pupil changes driven by between-trial differences in luminance, and attentional vigilance (Unsworth & Robison, 2015). This allowed us to examine task evoked pupillary responses (TEPRs) driven primarily by cognitive processes such as novelty detection and recognition memory (Goldinger & Papesh, 2012; Heaver & Hutton, 2011; Sirois & Jackson, 2012). In addition, because TEPRs are relatively slow, scores were averaged into 500ms bins for analysis. In this an all subsequently reported analyses, bins with more than 2 blinks were excluded, as blinks substantially alter derived pupil dilation measures.

TEPRs during the test array were then analyzed using a repeated measures ANOVA with time/bins (500ms, 1000ms, 1500ms, 2000ms) and response accuracy (correct, incorrect) as within subjects variables. Results revealed significant main effects for both time $F(3, 69)=20.978$, $p<.001$, $\eta^2_p=.477$, and response accuracy, $F(=1, 23)=8.640$, $p=.007$, $\eta^2_p=.273$ (Figure 11). These results suggest that although both conditions showed rapid returns to baseline over the course of the test array, pupils were significantly larger for incorrect trials than for correct trials. While the Bin x Accuracy interaction was not significant, additional follow-up simple effects test with a Sidak correction for multiple comparisons revealed that the difference between correct and incorrect TEPRs tended to increase over time, with significant differences at 500, 1500 and 2000ms (*p*=.036, .015 and .010, respectively).

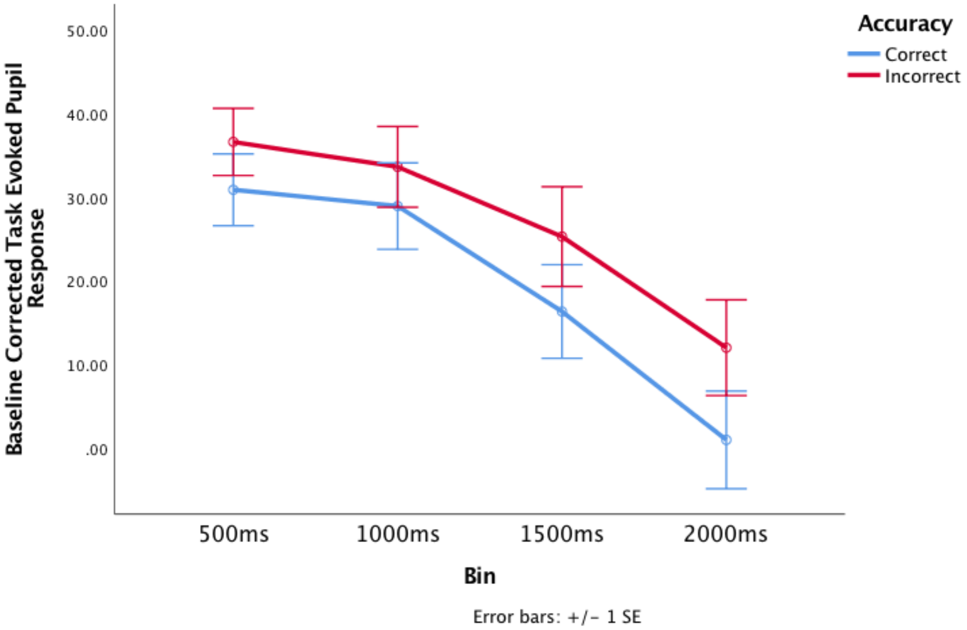


Figure 11: TEPRs as a Function of Response Accuracy and Bin.

Discussion

The goal of Experiment 1 was to identify pupil and eye movement correlates of visual STM performance (response accuracy) in adults. This was done via a direct replication of a task that has commonly been used as a way to quantify visual STM in adult participants with the added precision of eye tracking (Luck & Vogel, 1997). The results of this experiment have provided an important first step in learning about the similarities and differences between the memory systems that both infant and adults participants use to hold a subset of information active for a short period of time. Specifically, several low-level visual dynamics (e.g. saccade count, run count, average fixation duration, and TEPRs) were identified that were predictive of response accuracy and/or trial differences in adults. Adults show significant differences in the number of saccades during the test array (after the contents of the visual scene have been encoded into visual WM) as a function of response accuracy. On trials when adults are about to respond incorrectly, they make significantly more saccades. In addition to differences in saccade count, there are also significant differences in the average length of each fixation, with those about to respond incorrectly having shorter fixation duration (on average) compared to those who provide a correct response. In addition to these looking dynamics, there are also significant differences in changes in pupil size as a function of response accuracy.

Specifically, pupils get larger on trials when participants are about to respond incorrectly (i.e. there is less pupil change from baseline on correct trials).

Experiment 2 will utilize these low-level visual dynamics to look for evidence of memory maintenance in 5- and 11-month-old infants. Specifically, if visual behaviors and pupil responses during the test array differ systematically as a function of set size and change status, this would provide evidence that capacity estimates may be possible in infants. This will also provide a series of markers that are directly relatable to adult measures of visual STM, thus making infant/adult comparisons possible.

CHAPTER THREE

INFANT AND ADULT PASSIVE CHANGE-DETECTION TASK

Experiment 1 has successfully identified key visual dynamics (saccade count, average fixation duration, and TEPRs) that correspond to response accuracy in adults. If similar patterns hold while passively viewing arrays sizes likely to result in poor memory representations (e.g., larger set sizes, and change trials), then it may be possible to use visual dynamics to identify individual differences in visual STM performance, even in pre-verbal infants. Hence, the goal of Experiment 2A was to look for this evidence during a passive change-detection task in infants, and to determine if these visual dynamics are related to classic infant change detection measures like change preference. The goal of Experiment 2B was to ensure that relying on passive versions of the task did not fundamentally alter the visual dynamics related to adult behavioral accuracy. If these same visual dynamics present in Experiment 1 are present in both Experiment 2A and B, then it may be possible to establish a more precise, continuous measure of change detection. Further, if adults demonstrate similar visual behaviors in both Experiment 1 and 2B, then we can conclude that the addition of the overt behavioral response (button press) does not produce qualitatively different visual behaviors than passive viewing of a similar task. Together, these findings will help establish continuity between infants and adult paradigms and will determine if visual behaviors that accompany both passive and explicit change-detection are qualitatively similar.

EXPERIMENT 2A: DO INFANTS DEMONSTRATE SIMILAR MEMORY-DEPENDENT VISUAL DYNAMICS AS ADULTS?

Method

Participants

A total of 57 infants participated in a change-detection task. Six infants were removed due to fussiness and an additional 3 were excluded due to lack of interest. As a result, a total of 48 infants (21 females and 27 males) were tested at 5 (n=24) or 11 (n=24) months of age (+- 11 days). All infants were born within three weeks of their due date and had no reported birth defects or vision problems. Infant ethnicities were reported as follows: White (n=40), Black (n=3), Asian (n=1), Biracial (n=2) and 1 preferred not to answer. Infant names were obtained through the Tennessee Department of Health and Vital Statistics.

Apparatus

Infant Setup: A separate, but similar Eyelink 1000+ eye tracking system was used to collect data for the infant participants. Point of gaze data monocularly at 500 Hz and a 5-point dynamic calibration method was used. The camera was mounted on an Eyelink arm mount, which allowed the experimenter to adjust the position of the eye tracker without moving the infant and caregiver. A 16mm lens and 890nm infrared light emitter were used to measure continuous gaze. Both of these choices came as a recommendation from the SR-Research support team. The infant's left eye was tracked and pupil size was measured in terms of total area. Infants were tested in a dimly lit room and sat approximately 65cm in front of high-performance 24", 120Hz, Asus monitor (native resolution of 1920x1080), with a viewable surface of 45.5° (w) by 26.76° (h). Note, this standard infant setup was used for all infant tasks (Experiment 2A and Experiment 4).

Stimuli

Infants were shown multiple trials containing arrays consisting of 1, 2, or 3 of colored circles (see Figure 12). These arrays of colored circles were presented 8° from central fixation at 45°, 135°, 225°, and 315°. All circles measure 5° in diameter, and total eccentricity was 15.5° X 15.5°. Previous work (Ross-Sheehy & Eschman, 2019) suggests that infants do not prefer the changing item as set size 4, hence, for added power, this set size was excluded from the current experiment. Colors were randomly selected from a pool of 7 highly discriminable colors (red, purple, blue, white, black, green, and yellow; see Table 10 for RGB and luminance values). All arrays were presented on a 24" monitor with a viewable surface of 45.5° (w) by 26.76° (h), and all events were presented on a grey background. Each trial started with a fixation stimulus (multi-colored dynamic pinwheel). This was incorporated for two reasons. First, this acted as an attention-getter, ensuring that the infants were looking at the screen before each trial started. Second, it acted as "perceptual mask" that offered a clear

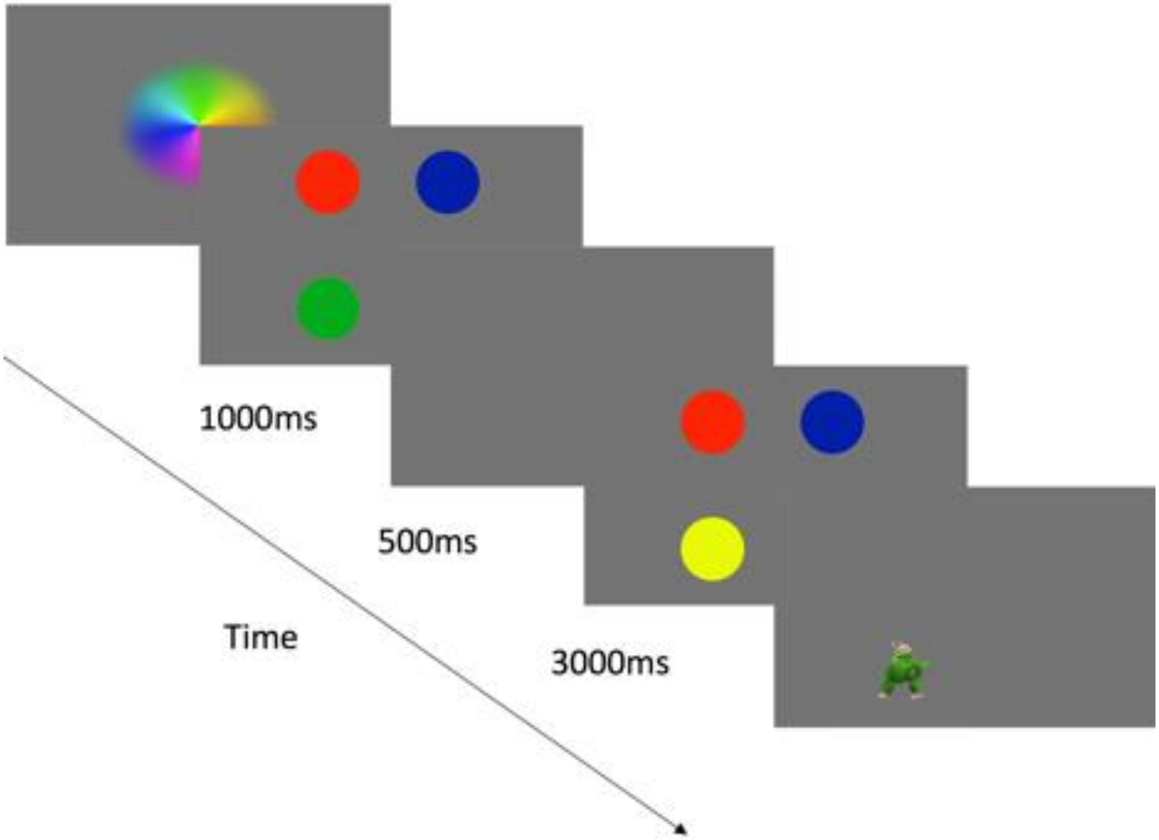


Figure 12: Experiment 2, Change Detection Task Set Size 3 Change Trial

demarcation between trials as a way to reduce the trial-to-trial interference. This ensured that the items were encoded in into visual STM, but not long-term memory (LTM) systems (Ross-Sheehy et al., 2003). This perceptual mask was followed by a *sample* array that was presented for 1000ms, to ensure sufficient time for encoding (long enough to go beyond sensory memory (Phillips, 1974) but short enough not to be encoded into LTM (Ross-sheehy et al., 2003)). Following a 500ms blank retention interval, infants were presented with a *test* array (3000ms). For half of the trials, the test array was exactly the same as the sample array (i.e., "same" trials). For the other half, the color of a single, randomly chosen circle changed (i.e., "different" trials). Immediately following the test array, infants saw a movie containing a musical dancing creature presented in the location of the color change (change trials only). This reward stimulus was designed to both increase general interest in the task and reinforce infant looking to the location of a color-change. Thus, if infants are able to notice a color change on a given test array, they should learn the contingency (i.e., color change = dancing creature) and increase looking to the changed item. Immediately after each trial, infants were presented with a multi-colored audiovisual attention getter/fixation stimulus that also served as a between-trial perceptual mask. This was designed to both attract infant attention to the center of the display in preparation for the next trial, and to purge the contents of visual STM from the previous trial.

Design and Procedure

This task incorporated a 2 x 3 design with both condition (change and no change) and set size (1, 2, and 3) manipulated within subject and age (5- and 11-month-old infants) as a between subject's variable. Participants viewed a maximum of 96 change-detection trials, but testing ended sooner if the infant became fussy, bored, or fatigued. On average, infants made it through 44 trials. A trained observer sat behind a black curtain and monitored eye tracking. Primary dependent measures included the proportion of looking to the change circle as a function of total looking, and total looking to the changing versus non-changing arrays as a function of set size. In addition, as in Experiment 1, eye tracking was used to assess pupillometry, saccades, fixations, and run count.

Results

Saccade Count

Saccade count was analyzed using a repeated measures ANOVA with change status (change, no-change) and set size (1, 2, 3) as within subjects variables and age (5- and 11-month-old infants) as between subject's variables. Results are summarized in Table 3, and revealed a significant main effect of set size $F(2,92)=55.998$, $p<.001$, $\eta^2_p=.549$. On average infants made the fewest number of saccades at set size 1 (1.98), compared to the most at set size 2 (2.253) and set size 3 (2.089). Follow-up simple effects tests with a Sidak correction for multiple comparisons show that there was a significant difference between set sizes 1 and 2, 2 and 3, and 1 and 3 (all *p*'s <.001) but there were no statistical differences between set size 2 and 3. This suggests that even infants will increase saccade count as set size increases, but only to a certain extent. Saccade count peaks at set size 2 and then dips at set size 3. This finding suggests that saccade count may be sensitive to set size differences. Additionally, there was a significant main effect of Age, $F(1, 46)=16.156$, $p<.001$, $\eta^2_p=.260$, as 5-month-old infants made more saccades than their 11-month-old counterparts. Results failed to produce a significant main effect of change status.

Run Count

Run count was also analyzed using a repeated measures ANOVA with change status (change, no-change) and set size (1, 2, 3) as within subjects variables and age (5- and 11-month-old infants) as a between subjects variable. Results revealed a significant main effect of set size, $F(2, 92)=109.242$, $p<.001$, $\eta^2_p=.704$, and change status, $F(1, 46)=242.8$, $p<.001$, $\eta^2_p=.841$. These results were qualified by a significant Set Size X Change Status X Age three-way interaction $F(2, 92)=5.321$, $p=.007$, $\eta^2_p=.104$ (Figure 13). Specifically, run count, which is the number of items that an individual fixates, is higher for infants on change trials compared to no-change trials. Based off of the adult data, individual infants make more eye movements and have longer runs when they respond incorrectly to similar change-detection tasks. Thus, these results suggest that

Table 3: Analysis of Variance Visual Dynamics Results (Experiment 2A)

Infant	***F***	***p***	***df***	df_{error}	η^2
SACCADE COUNT					
Set Size	55.998	***<.001***	2	92	0.549
Age	16.156	***<.001***	1	46	0.26
Change Status	0.016	0.901	1	46	<.001
Set Size X Age	0.623	0.539	2	92	0.539
Change Status X Age	0.429	0.516	1	46	0.009
Set Size X Change Status	1.724	0.184	2	92	0.036
Set Size X Change Status X Age	2.821	0.065	2	92	0.058
RUN COUNT					
Set Size	109.242	***<.001***	2	92	0.704
Age	0.462	0.5	1	46	0.01
Change Status	242.8	***<.001***	1	46	0.841
Set Size X Age	1.431	0.244	2	92	0.03
Change Status X Age	0.134	0.716	1	46	0.003
Set Size X Change Status	4.626	***0.012***	2	92	0.091
Set Size X Change Status X Age	5.321	***0.008***	2	92	0.104
AVERAGE FIXATION DURATION					
Set Size	34.946	***<.001***	2	92	0.432
Age	12.481	***0.001***	1	46	0.213
Change Status	0.066	0.799	1	46	0.001
Set Size X Age	2.238	0.112	2	92	0.432
Change Status X Age	0.689	0.411	1	46	0.015
Set Size X Change Status	0.542	0.584	2	92	0.012
Set Size X Change Status X Age	0.023	0.977	2	92	<.001
TEPRs					
Time (bins)	1.388	0.231	5	185	0.036
Age	1.547	0.221	1	37	0.228
Set Size	8.659	***<.001***	1	37	0.19
Change Status	0.229	0.635	1	37	0.006
Change Status X Age	1.267	0.268	1	37	0.033
Bin X Set Size	1.916	***0.042***	10	370	0.049
Bin X Set Size X Age	1.914	0.114	10	370	0.049
Bin X Change Status	3.041	***0.012***	5	185	0.076
Bin X Change Status X Age	1.764	0.122	5	185	0.046
Set Size X Change Status	0.891	0.414	2	74	0.024
Set Size X Change Status X Age	0.314	0.731	2	74	0.008
Bin X Set Size X Change Status	0.407	0.943	10	370	0.011
Bin X Set Size X Change Status X Age	0.448	0.922	10	370	0.012

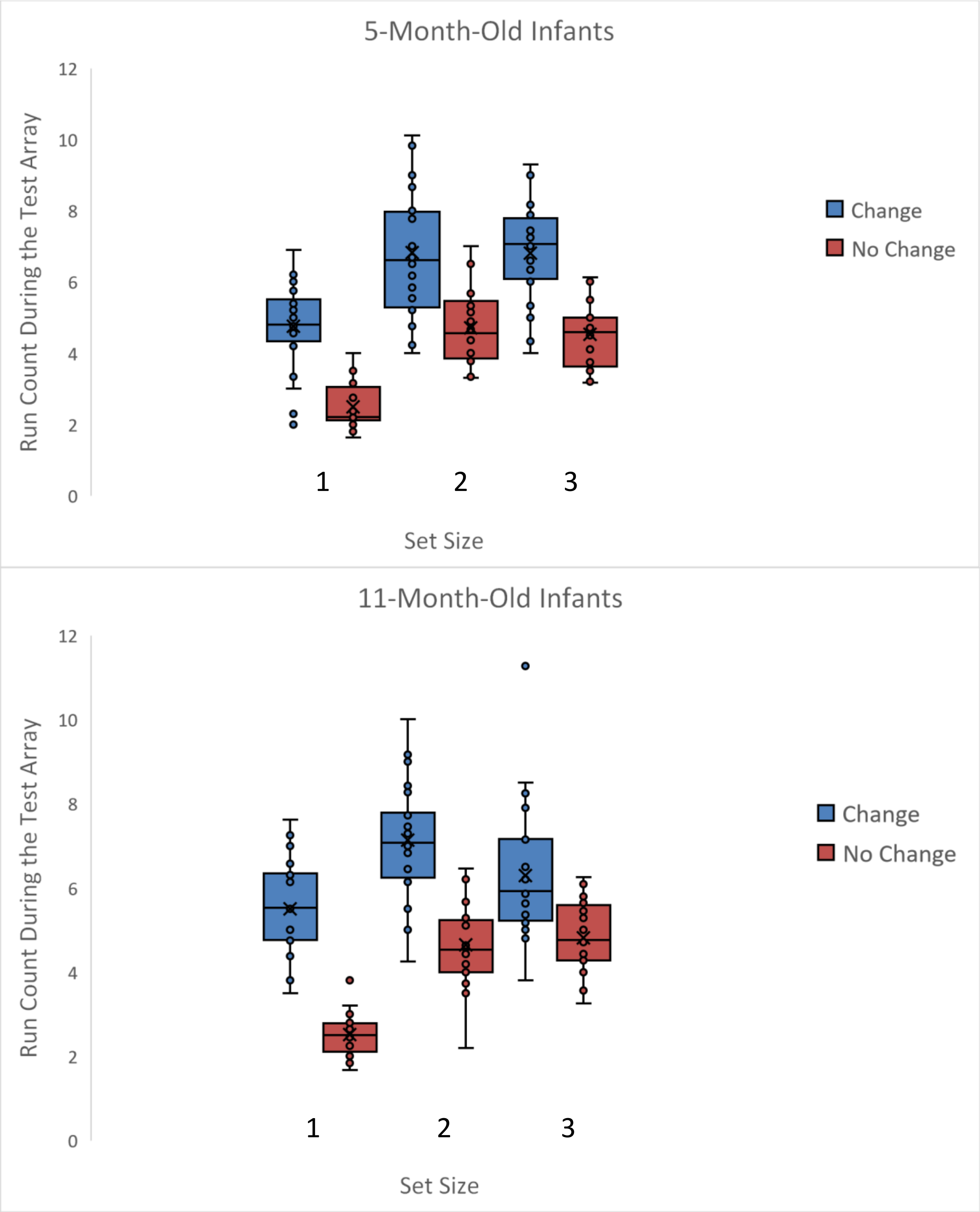


Figure 13: Infant Run Count as a Function of Age, Change Status, and Set Size

even infants were sensitive to change status, as their runs are considerably longer when something changes.

Average Fixation Duration

Average fixation duration was analyzed using a repeated measures ANOVA with change status (change, no-change) and set size (1, 2, 3) as within subjects variables and age (5- and 11-month-old infants) as between subject's variables. There was a significant effect of set size $F(2, 92)=34.946$, $p<.001$, $\eta^2_p=.432$. Similar to both saccade count and run count, there was no main effect of age. Infants had the longest average fixation durations at set size 1 (344.29ms) and the shortest at set size 2 (310.825ms). Set size 3 fell in the middle (334.1ms). There was also a significant main effect of age, $F(1, 46)=12.481$, $p=.001$, $\eta^2_p=.213$, in that 11-month-old infants had longer average fixation durations than their 5-month-old counter parts. There was no main effect of change status ($p=.799$). This could be due to the fact that infants had shorter average fixation (436ms) durations compared to adults (498ms).

Change Preference

Change preference, or the proportion of time participants spend looking to the item that changed compared to the rest of the non-changed items was also computed. Change preference was compared to chance at set size 2 (.5) and set size 3 (.333) using a one-sample *t*-test. In addition, to ensure that the results were a true measure of preference for the changing item, only change trials where the participants made at least one look to the changed item and at least one look to a non-changed item were analyzed. Change preference at each set size was then compared to chance using a one-sample *t*-test. At 5-months of age, infants did not show a significant preference for the changed item during the test array, for set size 2 (*p*=.725), but did show a significant preference for the changed item for set size 3 arrays, *t*(23)=3.5, *p*=.002. Additionally, 11-month-old infants failed to display a significant preference for the changed item during the test array at set size 2 (*p*=.217) but, like the 5-month-old infants, show a significant preference for the changed item at set size 3, *t*(23)=5.729, *p*<.001. While these results may seem puzzling, there are several likely contributors to this pattern of performance. First, as seen in Ross-Sheehy and Eschman (2019) change preference may be driven primarily by recency effects rather than memory for the entire array and may be quite noisy. Second, our criterion that infants look at both the change and at least one no-change circle effectively reduced power, making it difficult to demonstrate a change preference. Both of these possibilities suggest these results should be interpreted with caution. There is, however, a third possibility. Perhaps in set size 2 arrays, the attention-getting properties of both the change circle (i.e., novelty preference) and the no-change circle (i.e., familiarity preference) effectively cancel each other out, resulting in a null effect. This could potentially explain why there is no effect for set size 2, but a strong effect for set size 3. One way to test this possibility, is by examining pupil responses during the test array. If novelty and familiarity detection are influencing eye movements, we should see distinctly different TEPRs while fixating novel versus familiar circles. Although this analysis is not possible with our current data, future work will be conducted to assess this possibility.

Pupil Analyses

Just like Experiment 1, TEPRs were chopped into 500ms bins and baseline corrected (first 100ms of the sample array). Due to the fact that pupil dilation is estimated using reflected infrared light, several factors unique to infant participants may at times lead to the increased occurrence of “pupil blow out”, or the inclusion of the entire iris and sclera in the pupil diameter estimate (e.g., decreased pigment in retina, use of 16mm lens rather than 25 or 35mm lens, increased head movement, etc). These errors are easy to detect, and lead to pupil estimates several orders of magnitude larger than typical estimates. To address this, an outlier analyses with a 1.5 interquartile range (IQR) was conducted. This analysis identified 3 individuals as outliers, and these individuals were subsequently dropped from this analysis. A repeated measures ANOVA time (500ms, 1000ms, 1500ms, 2000ms, 2500ms, 3000ms), set size (1-3), change status (change, no change) as within subjects variables and age (5mo,11mo) as a between subjects variable demonstrated a significant main effect of set size, F(5, 74) = 8.659, $p<.001$, $\eta^2_p=.190$, which was qualified by a significant set size by bin interaction, $F(10, 370) = 1.916$, $p=.042$, $\eta^2_p=.049$ (Figure 14). This finding suggests that even in infants, TEPRs are sensitive to stimulus properties like visual load. It is likely that infants perform best at the smallest set sizes and this just so happens to be where we see the largest change from baseline in pupil size. This analysis also revealed a significant bin by change status interaction, $F(5, 185) = 3.041$, $p=.012$, $\eta^2_p=.076$ (Figure 15), with greater pupil change from baseline for change trials relative to no-change trials. This is important and suggests that TEPRs may also be sensitive to novelty detection, resulting in relatively large deviations away from baseline, followed by a rapid return to baseline.

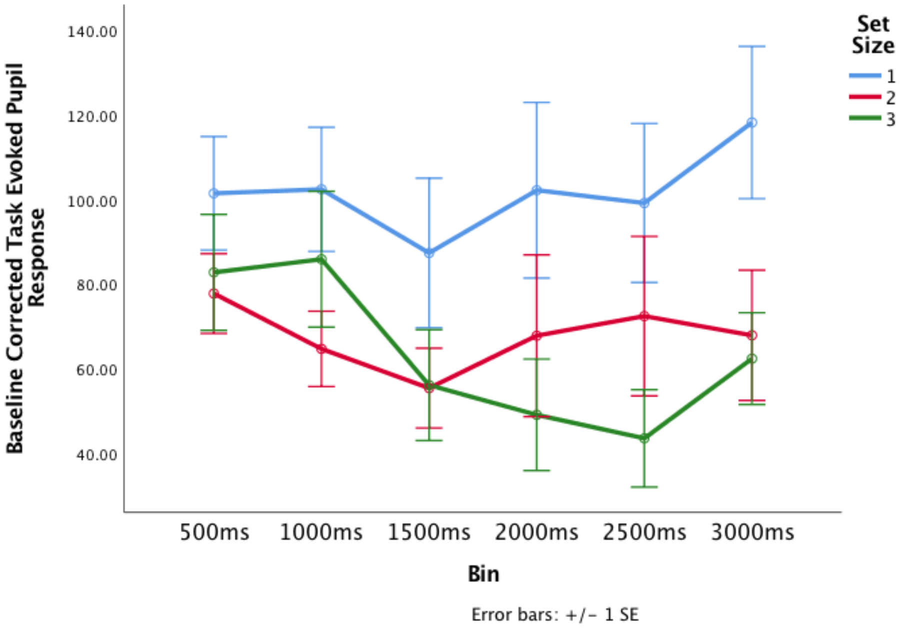


Figure 14: Infant TEPRs as a Function of Bin and Set Size

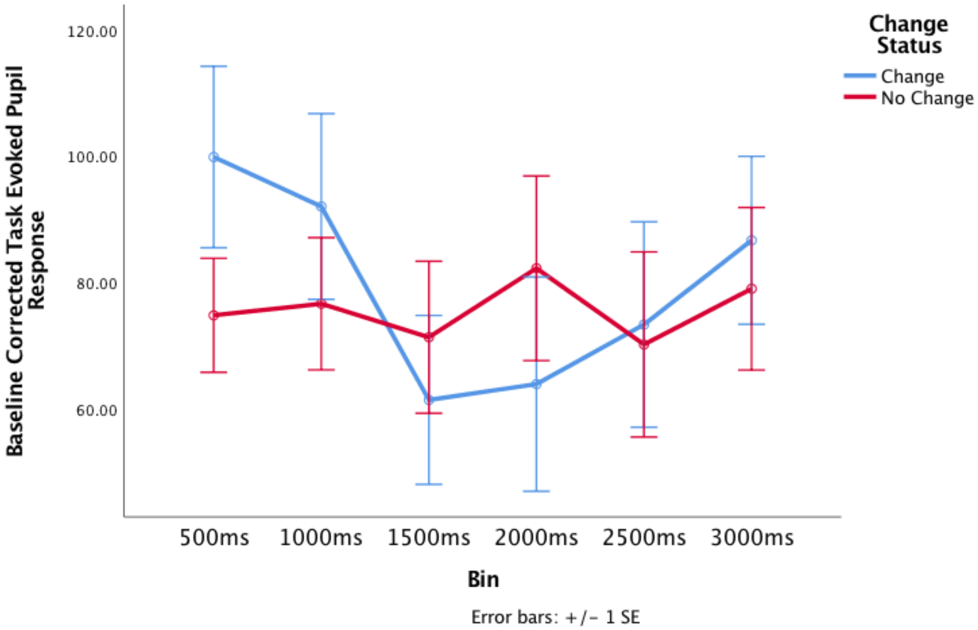


Figure 15: Infant TEPRs as a Function of Bin and Change Status

Discussion

The goal of Experiment 2A was to look for the visual dynamics identified in Experiment 1 using a passive change-detection task. Results from Experiment 2A suggest that infants are engaging in qualitatively similar visual behaviors as their adult counterparts. Specifically, just like adults, the number of saccades, run count, and average fixation duration differ with set size. It is assumed that lower set sizes are more likely to produce more accurate responses so demonstrating that these visual dynamics scale with set size is an important first step in demonstrating that these dynamics (related to response accuracy in adults) are similar across both infant and adult participants. Additionally, run count and TERPs differed significantly as a function of change status. This is incredibly important, suggesting that these visual dynamics in infants are sensitive to change status. This is a vital second step in demonstrating that visual dynamics are similar across both infant and adult participants.

While these data strongly suggest that infants are engaging in similar visual behavior to adult participants in this modified change detection task, it is possible that because adults are providing an overt behavioral (button) response, that this in some way completely changes that task itself. Therefore, in order to know if visual dynamics (the very same ones that are related to accuracy) are 1) the same for infants and adults, and 2) the same for passive/active tasks, it is imperative to have adult participate in a modified change-detection task, without any kind of overt behavioral response. Experiment 2B aims to achieve these goals by having adults participate in the exact same experiment used with the infants.

EXPERIMENT 2B: DO ADULTS DEMONSTRATE SIMILAR MEMORY-DEPENDENT VISUAL DYNAMICS IN A PASSIVE CHANGE-DETECTION TASK?

Method

Participants

A total of 26 adults participated in a passive change-detection task identical to Experiment 2A. Two participants' data was removed from the final data set due to difficulty during calibration that resulted in poor eye tracking data quality. The final sample consisted of 24 adults (13 females and 11 males, *M* age = 19.15 years). Ethnicities were reported as follows: White (n=21), Black (n=1), Pacific Islander (n=1) and Asian (n=1). All participants had normal or corrected to normal acuity, and were screened for colorblindness (Ishihara, 1960). Additionally, all participants reported no serious head injuries or neurological disorders. Adult participants were recruited from the University of Tennessee's online recruitment tool and were awarded course credit for their participation.

Stimuli and Apparatus

The eye tracking setup was identical to Experiment 1, and stimuli and procedure were identical to Experiment 2A.

Results

Saccade Count

Saccade count was analyzed using a repeated measures ANOVA with change status (change, no-change) and set size (1, 2, 3) as within subjects variables. Results are summarized in Table 4. Similar to Experiments 1 and 2A, there was a significant main effect of set size $F(2, 46)=29.701$, $p<.001$, $\eta^2_p=.564$, demonstrating significantly more saccades for larger set sizes (set size 1 = 4.022, set size 2 = 4.999, set size 3 = 5.469). While the neither the main effect of change status nor the change status by set size interaction were significant, it should be noted that all three of these set sizes are well within typical adults visual STM capacity; therefore, it is likely that the lack of significant effects was partially due to a ceiling effect.

Run Count

Run count was also analyzed using a repeated measures ANOVA with change status (change, no-change) and set size (1, 2, 3) as within subjects variables. There was also a main effect of set size $F(2, 46)=58.027$, $p<.001$, $\eta^2_p=.716$, demonstrating significantly longer runs on trials with larger set sizes (Figure 16). There was also a main effect of change status $F(1, 23)=271.132$, $p<.001$, $\eta^2_p=.922$. Participants had longer runs on change trials compared to the no change trials. This important finding suggests that even in the absence of an

Table 4: Analysis of Variance Visual Dynamics Results (Experiment 2B)

Adults	***F***	***p***	***df***	***df***	***η*²**
SACCADE COUNT					
Set Size	29.701	***<.001***	2	46	0.564
Change Status	2.861	0.104	1	23	0.111
Set Size X Change Status	1.299	0.283	1	23	0.053
RUN COUNT					
Set Size	58.027	***<.001***	2	46	0.716
Change Status	271.132	***<.001***	1	23	0.922
Set Size X Change Status	3.967	***0.026***	2	46	0.147
AVERAGE FIXATION DURATION					
Set Size	13.484	***<.001***	2	46	0.37
Change Status	0.19	0.667	1	23	0.008
Set Size X Change Status	3.863	***0.028***	2	46	0.144
TEPRs					
Time (bins)	10.288	***<.001***	5	105	0.329
Set Size	4.682	***0.015***	2	42	0.182
Change Status	0.156	0.697	1	21	0.007
Bin X Set Size	5.055	***<.001***	10	210	0.194
Bin X Change Status	0.849	0.518	5	105	0.039
Set Size X Change Status	2.515	0.093	2	42	0.107
Bin X Set Size X Change Status	2.638	***0.005***	10	210	0.112

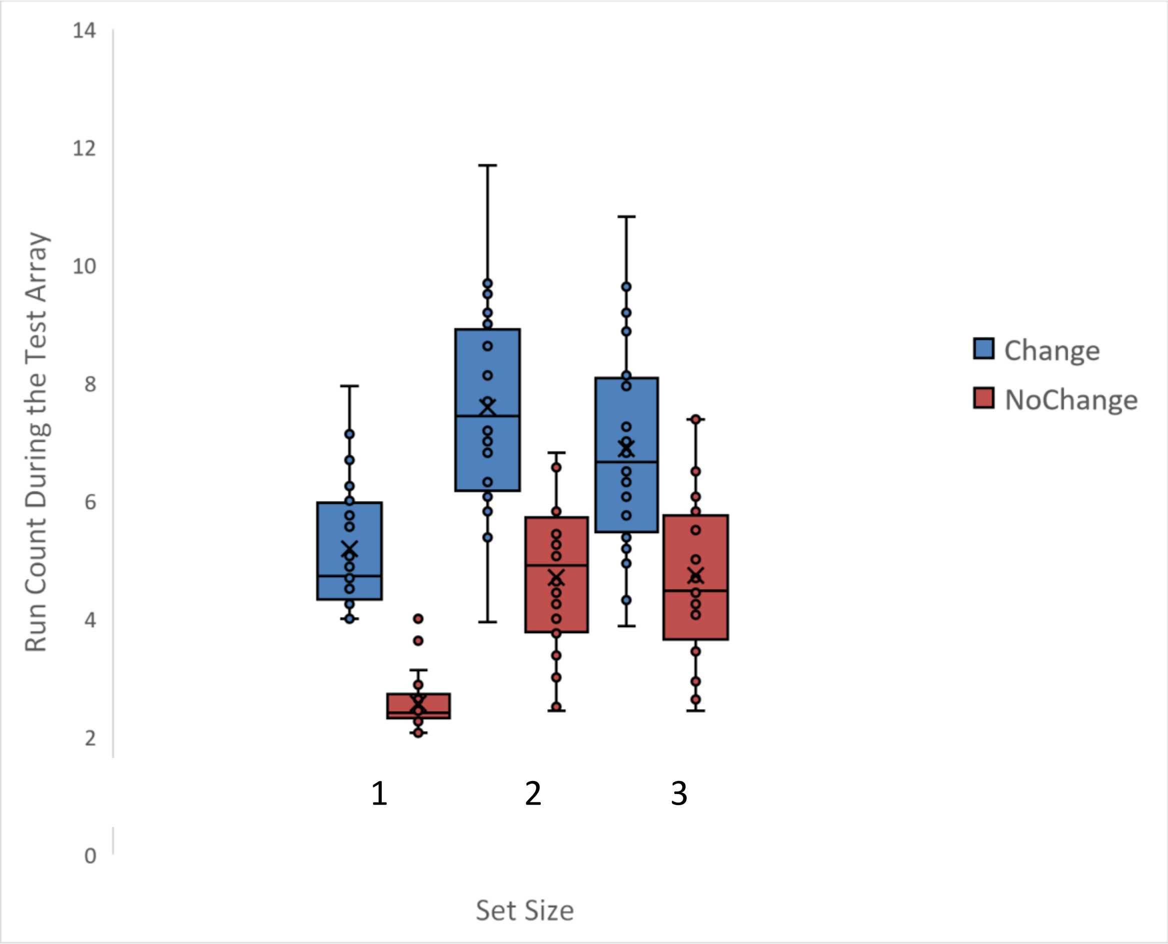


Figure 16: Adult Run Count as a Function of Change Status and Set Size

overt behavioral response, it is still possible to detect significant differences in visual behavior as a function of trial type. These results were qualified by a significant set size by change status interaction $F(2, 46)=3.967$, $p<.001$, $\eta^2_p=.683$ revealing that differences between run count for the change and no change trials decrease as set size increases. Sidak follow-up tests show that there was a significant difference between change and no change trials at all set sizes (all *p*'s <.001).

Average Fixation Duration

Average fixation duration was analyzed using a repeated measures ANOVA with change status (change, no-change) and set size (1, 2, 3) as within subjects variables. There was a significant effect of set size $F(2, 46)=13.484$, $p<.001$, $\eta^2_p=.370$, revealing longer average fixation durations for smaller set sizes. In addition, there was also a significant set size by change status interaction $F(2, 46)=3.863$, $p=.028$, $\eta^2_p=.144$ (Figure 17). Sidak follow-up tests show that there was a marginally significant difference between change and no change trials at set size 1 only ($p=.064$).

Pupil Analyses

As in Experiment 1 and 2A, TEPRs were divided into 500ms bins and baseline corrected (first 100ms of the display array). A repeated measures ANOVA with time (500ms, 1000ms, 1500ms, 2000ms, 2500ms, 3000ms), set

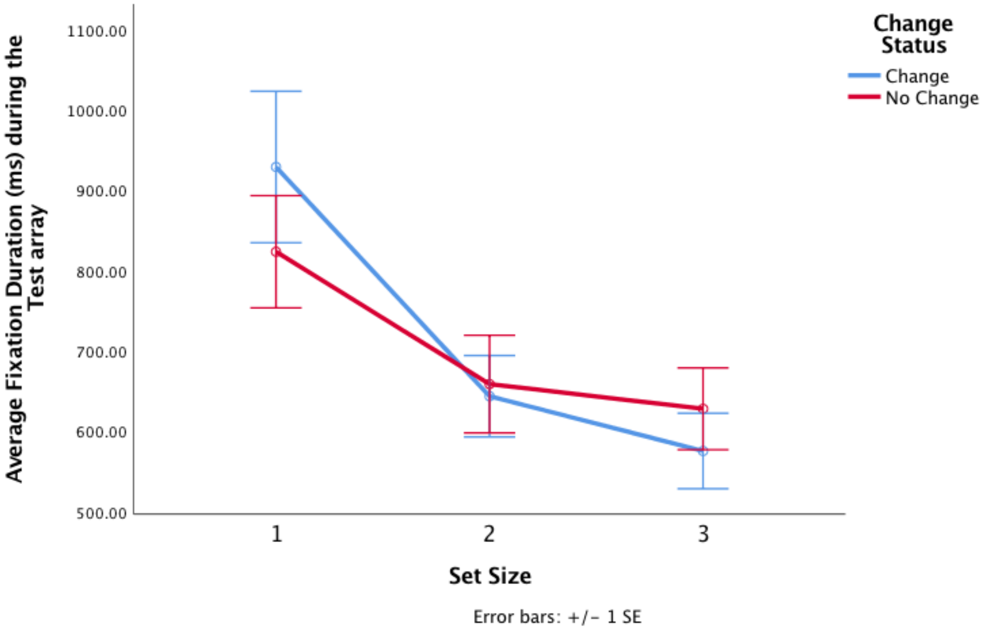


Figure 17: Adult Average Fixation Duration as a Function of Change Status and Set Size

size (1-3) and change status (change, no-change) as within subjects variables revealed a main effect of bin, F(5, 105) = 10.288, *p*<.001, η^2_p=.329 (Figure18). There was also a main effect of set size F(2, 142) = 4.682, *p*=.015, η^2_p=.182 (Figure 19), and a significant 3-way interaction (Bin X Set Size X Change Status) F(10, 210) = 2.638, *p*=.005, η^2_p=.112. These results were consistent with the infant data in Experiment 2A in that pupils were largest for the smallest set size and got smaller as set size increased.

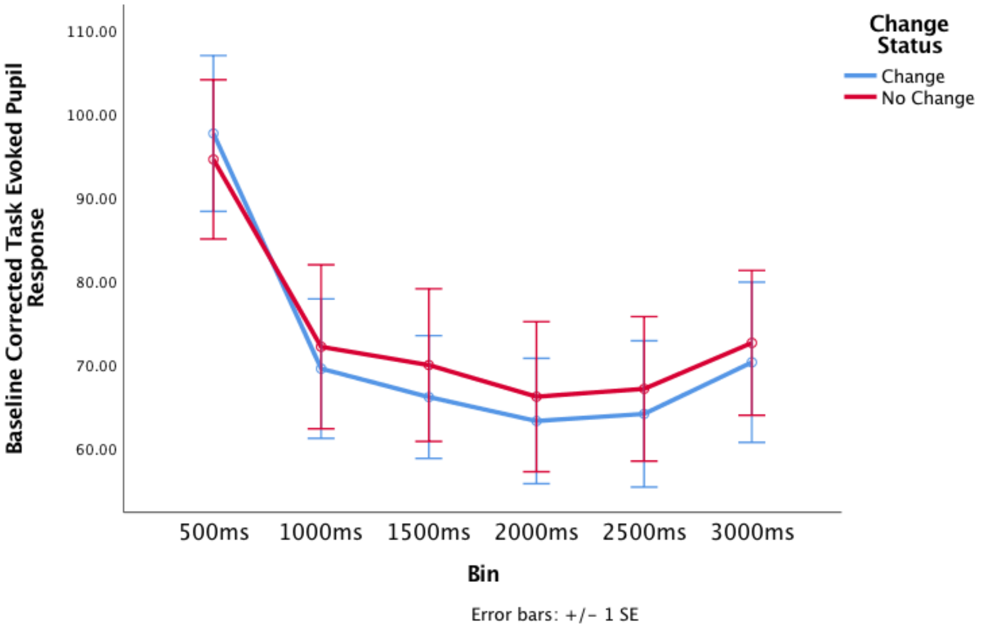


Figure 18: Adult TEPRs as a Function of Bin and Change Status

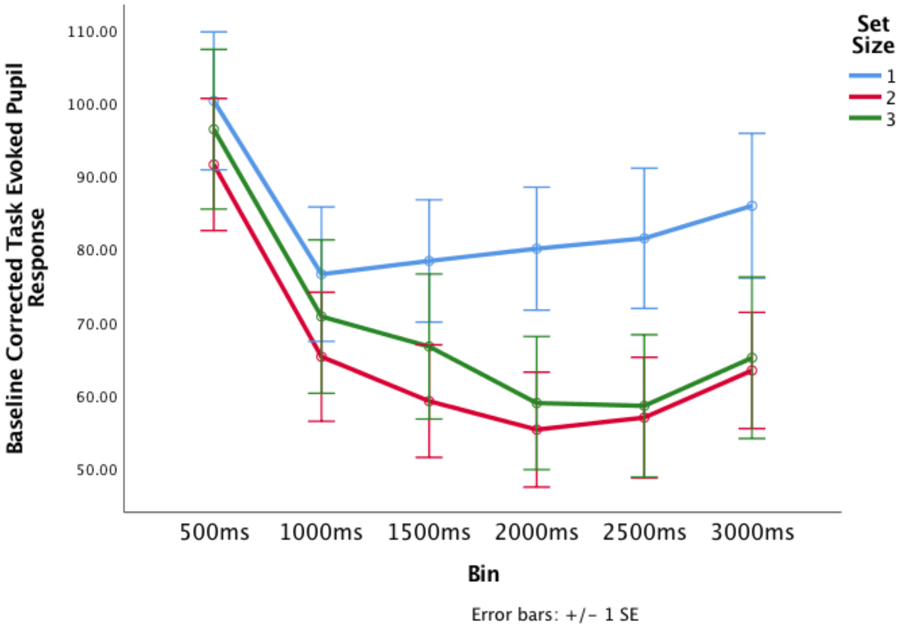


Figure 19: Adult TEPRs as a Function of Bin and Set Size

Discussion

The adult data from Experiment 2B are strikingly similar to the infant data in Experiment 2A. As was the case with infants, adult saccade count, run count, and average fixation durations all changed as a function of set size. In addition to these set size differences, there were also significant differences in run count as a function of change status. Again, just like the infants, the adults made significantly more runs on trials when one of the items changed, compared to the "no-change" trials. Additionally, the pupil data (in terms of set size effects) were also similar, suggesting a qualitatively similar cognitive processes is being measured in both age groups. The effects of change status may seem different than the infants but recall the largest set size used in this task was 3, which is within an adult's typical visual STM capacity—likely resulting in a ceiling effect. Although it is currently not possible to know whether the adult has successfully noticed the changed item, larger set sizes incur the greatest visual STM processing load. Recall that Experiment 1 demonstrated a relationship between visual dynamics and response accuracy. Experiment 2 has subsequently established a similar pattern while participants were passively viewing array sizes that are likely to result in poor memory representations (e.g., larger set sizes, and change trials). As a result, it is reasonable to suggest that it may be possible to use visual dynamics to identify individual differences in visual STM capacity, even in pre-verbal infants

While these results are promising, it is worth noting that several of the task parameters are different in Experiments 1 and 2B. Specifically, in addition to the lack of behavioral button response, the encoding period is longer, the set sizes were smaller, the size of the array items were larger, and the total eccentricity was larger. Thus, to bridge the gap between infant and adult paradigms, Experiment 3 will combine the larger set sizes and button-press responses from Experiment 1 with several infant parameters from Experiment 2B. In addition to increasing power for the accuracy measure, these manipulations will help determine if typical infant parameters including large total eccentricates and longer test durations influence visual STM.

CHAPTER FOUR

EXPERIMENT 3: DO INFANT SPECIFIC ADAPTATIONS IN CHANGE-DETECTION TASKS INFLUENCE ADULT VISUAL BEHAVIOR?

The goal of Experiment 3 was to test adults in a change-detection task with parameters more typical of those commonly used in infant change-detection tasks—in an attempt to determine if typical infant adaptations (i.e., larger array items, longer encoding durations) are influencing STM performance. In other words, this experiment was designed to see if the results from Experiment 2, were merely a byproduct of the task parameters. Recall that the individual items used in Experiment 1 were relatively small (0.65°) and arrays were constrained to appear within a 9.8° X 7.3° region of the screen. If increasing the size of the individual items and the total eccentricity forces adults to make more eye movements, this may be an indication that these visual dynamics are influenced by the task parameters themselves and not purely a byproduct of visual STM performance. Additionally, encoding time was significantly shorter than is typical in infant change-detection paradigms (100ms compared to 1000ms), giving adults less time to demonstrate these differences. Thus, Experiment 3 bridges the gap between classic adult change-detection tasks (Luck & Vogel, 1997) and infant change-detection tasks. To accomplish this, adults were tested in a *hybrid* change-detection task. As in Experiment 1, set size and change status were manipulated to determine each individual's visual STM capacity. However, unlike

Experiment 1, Experiment 3 included relatively long encoding durations, larger array items, and larger total array eccentricity. These changes also afforded comparisons between adult visual behavior during encoding and subsequent behavioral accuracy. These comparisons were impossible for Experiment 1, as encoding durations were very brief (100ms) and array total eccentricity was very small (9.8° X 7.3). Additionally, if adult results reveal qualitatively similar capacity effects as in Experiment 1 and previous work (e.g., Luck and Vogel, 1997), this would provide more evidence that the infant task is in fact tapping a similar STM system that is being measured in adults.

Method

Participants

A total of 27 adults participated in a modified change-detection task. Two participants' data were removed from the final data set due to difficulty during calibration that resulted in poor eye tracking data quality (lost gaze on more than 50% of the sample), and one participant's data was removed due to experimenter error. The final sample consisted of 24 adults (19 females, 4 males, and one participant declined to answer, *M* age = 19.04 years). Ethnicities were reported as follows: White (n=20), Asian (n=2), and two or more races (n=2). All participants had normal or corrected to normal acuity and were screened for colorblindness (Ishihara, 1960). Additionally, all participants reported no serious head injuries or neurological disorders. Adult participants were recruited from the University of Tennessee’s online recruitment tool and were awarded course credit for their participation.

Stimuli and Apparatus

The eye tracking setup was identical to the standard adult eye tracking set-up outlined in Experiment 1. Adults participated in a modified change-detection task to assess visual STM capacity (see Figure 20). Consistent with Experiments 1 and 2, participants were shown multiple trials containing arrays consisting of colored circles. Set sizes consisted of 3, 6, or 9 items. The colors of the circles were drawn randomly without replacement from a pool of 10 discriminable colors (see Table 10). Note that unlike Experiment 1, 10 colors were used to ensure that there would never be two of the same color in the same array. Encoding time was also increased to reflect typical infant encoding durations (e.g., 1000ms for Experiment 2). After the 3000ms test array, participants were prompted to respond via button press if the second array was the same or different from the first array.

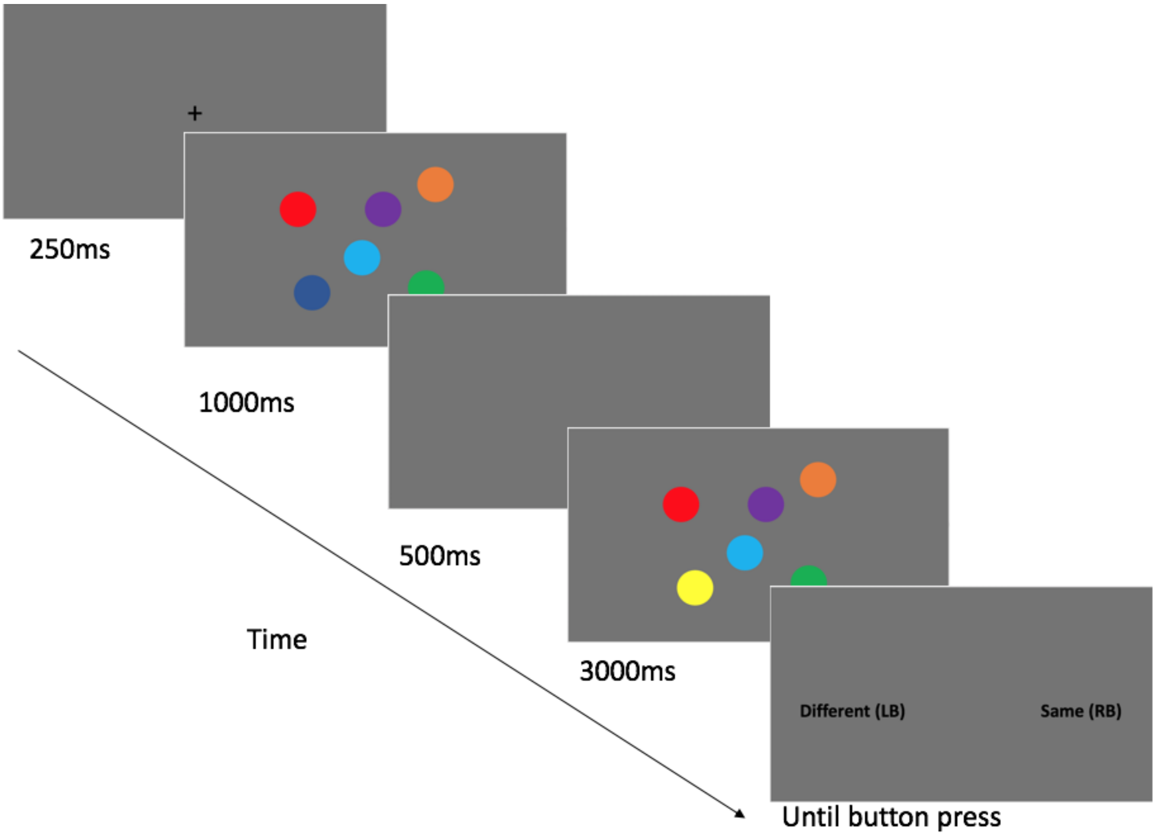


Figure 20: Experiment 3, Change Detection Task Set Size 6 Change Trial

Design and Procedure

This task incorporated a 2 x 3 design with both condition (change and no change) and set size (3, 6, and 9) manipulated within subject. Each individual viewed up to 40 blocks of trials, and each block contained one of every possible trial type, presented randomly within block (240 trials total). The experimenter was seated behind a black curtain and monitored eye movements from a small video monitor. Adults were instructed to fixate a small black cross, located in the middle of the screen, to initiate the trial presentation (the trial started after they fixated the cross for 250ms). Stimuli consisted of arrays of colored circles, presented in a random location within a total eccentricity was 15.5° X 15.5°. All circles measured 5° in diameter. Once the participant fixated the fixation cross, the trial presentation automatically began (1000ms memory array, 500ms retention interval, 3000ms test array).

Results

Percent Correct

Just as in Experiment 1, performance was the highest at the smallest set size (set size 3) with increasing errors as set sizes increased (Figure 21). Results of a repeated measures ANOVA with change status (change, no-change) and set size (3, 6, 9) as within subjects variables revealed a significant main effect of set size $F(2,46)=133.88$, $p<.001$, $\eta^2_p=.853$ (Figure 21). Follow-up simple effects tests with a Sidak correction for multiple comparisons show that percent correct at set size 3 is significantly higher than at set size 6 ($p<.001$), and that percent correct at set size 6 is significantly higher that it is at set size 9 ($p<.001$). In addition to the main effect of set size, there was also a main effect of change status $F(1,23)=24.905$, $p<.001$, $\eta^2_p=.520$, in that participants had a higher percent correct on no-change trials.

Saccade Count

Saccade count was analyzed using a repeated measures ANOVA with change status (change, no-change), set size (3, 6, 9), and response accuracy (correct, incorrect) as within subjects variables. Results are summarized in Table 5. As in Experiment 1, results reveal a significant main effect of accuracy, $F(1,16)=8.954$, $p=.015$, $\eta^2_p=.319$, with higher saccade counts for incorrect trials (Figure 22). Importantly, this result was also found in Experiment 1, and demonstrates that adopting longer encoding durations, larger array items, and larger total eccentricity does not disrupt the relation between saccade count and response accuracy in adult participants. There was no main effect of change status or set size.

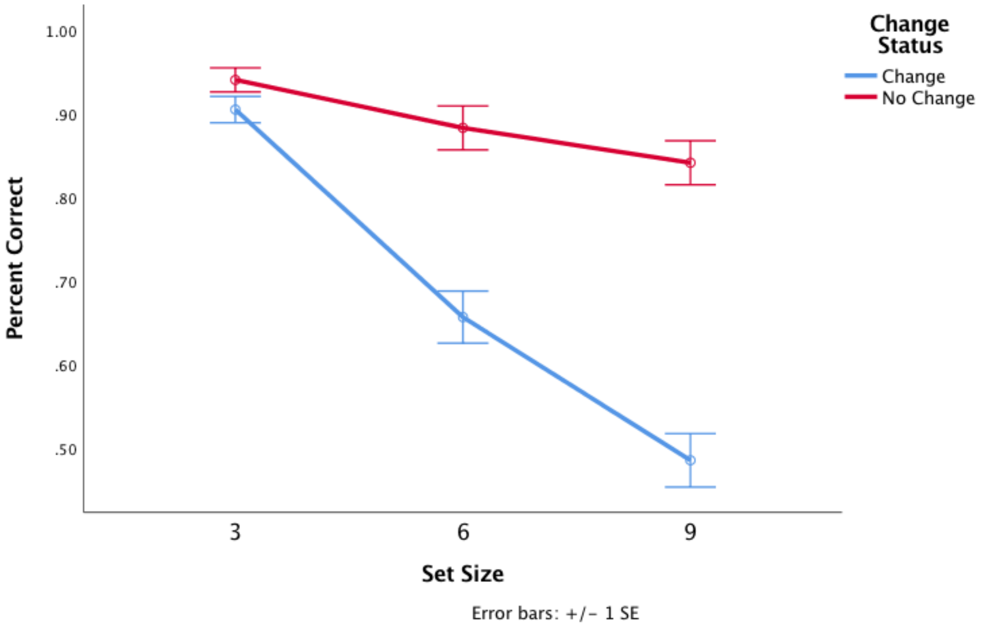


Figure 21: Average Percent Correct as a Function of Change Status and Set Size

Table 5: Analysis of Variance Visual Dynamics Results (Experiment 3)

	F	*p*	*df*	*df*error	η^2
SACCADE COUNT					
Set Size	1.725	0.194	2	32	0.097
Change Status	1.347	0.263	1	16	0.078
Accuracy	7.493	***0.015***	1	16	0.319
Set Size X Change Status	0.496	0.614	2	32	0.03
Set Size X Accuracy	1.11	0.342	2	32	0.065
Change Status X Accuracy	0.052	0.822	1	16	0.003
Set Size X Change Status X Accuracy	1.428	0.255	2	32	0.082
RUN COUNT					
Set Size	12.48	***<.001***	2	32	0.438
Change Status	1.888	0.188	1	16	0.106
Accuracy	5.993	***0.026***	1	16	0.272
Set Size X Change Status	1.825	0.178	2	32	0.102
Set Size X Accuracy	1.046	0.363	2	32	0.061
Change Status X Accuracy	0.138	0.715	1	16	0.009
Set Size X Change Status X Accuracy	1.652	0.208	2	32	0.094
AVERAGE FIXATION DURATION					
Set Size	1.562	0.225	2	32	0.089
Change Status	0.64	0.435	1	16	0.038
Accuracy	3.696	0.073	1	16	0.188
Set Size X Change Status	0.161	0.852	2	32	0.01
Set Size X Accuracy	0.302	0.741	2	32	0.019
Change Status X Accuracy	1.347	0.263	1	16	0.078
Set Size X Change Status X Accuracy	0.016	0.984	2	32	0.001
TEPRs					
Time (bins)	2.874	***0.018***	5	115	0.111
Accuracy	3.725	0.066	1	23	0.139
Bin X Accuracy	2.709	***0.024***	5	115	0.105

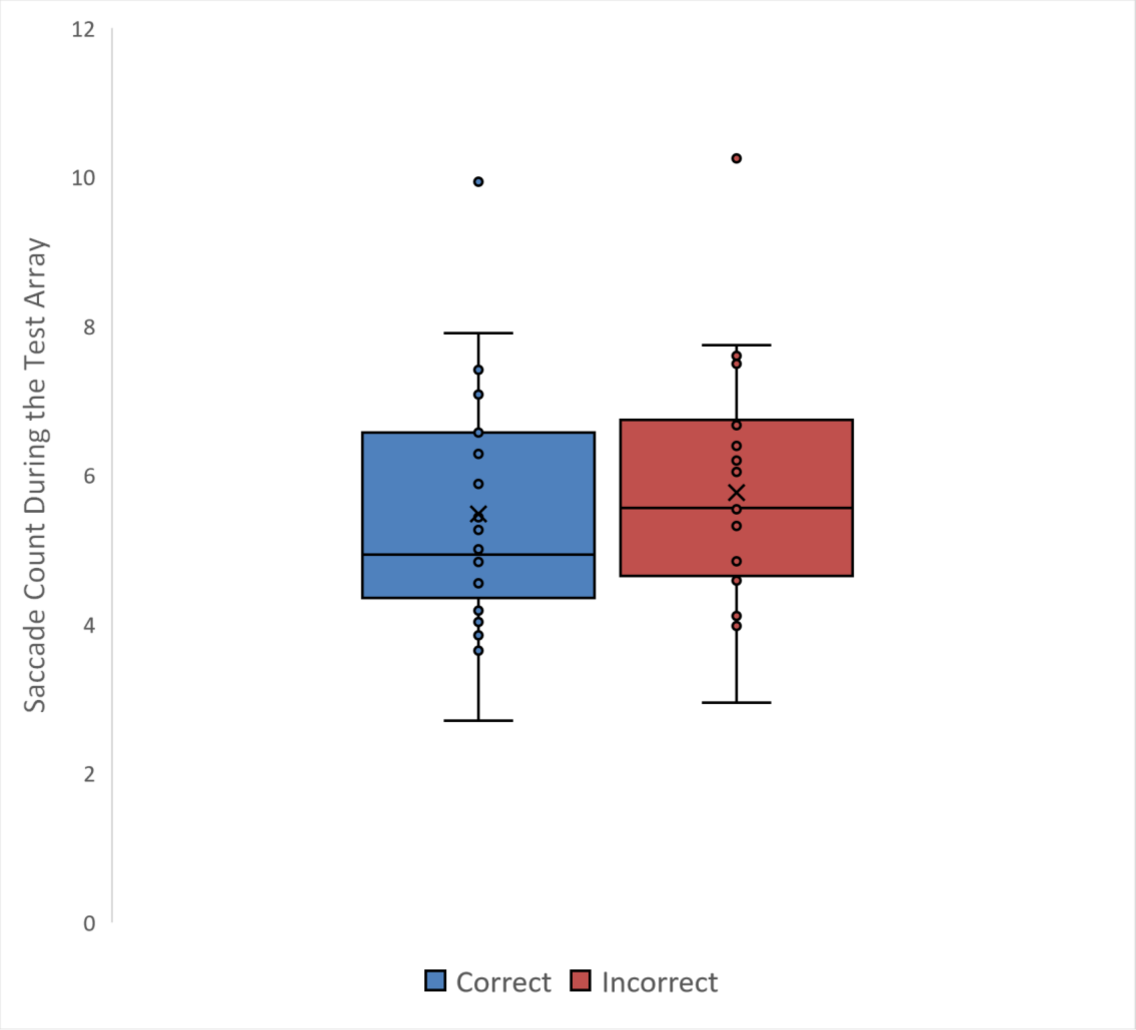


Figure 22: Saccade Count During the Test Array as a Function of Response Accuracy

Run Count

Run count was analyzed using a repeated measures ANOVA with change status (change, no-change), set size (3, 6, 9), and response accuracy (correct, incorrect) as within subjects variables. Results reveal a significant main effect of set size $F(2,32)=12.48$, $p<.001$, $\eta^2_p=.438$, with longer run counts on trials with larger set sizes. There was also a significant main effect of accuracy, $F(1,16)=5.993$, $p=.026$, $\eta^2_p=.272$, with significantly longer run counts for incorrect trials (Figure 23). This suggests that adults were fixating more items, multiple times when they responded incorrectly.

Average Fixation Duration

Just like Experiment 1, average fixation duration was calculated by computing the average duration (in milliseconds) of each individual fixation during the test array. Again, data were analyzed via a repeated measures ANOVA with change status (change, no-change), set size (4, 8, 12) and response accuracy (correct, incorrect) as within subjects variables. Results failed to reveal a significant main effect of set size ($p=.225$). They did reveal a marginal main effect of response accuracy $F(1,16)=3.696$, $p=.073$, $\eta^2_p=.439$ (Figure 24). Again, just like Experiment 1, this is important and suggests that longer fixations are either directly contributing to higher accuracy, or that individuals who tend to fixate longer, tend to have higher WM ability. Taken together with saccade

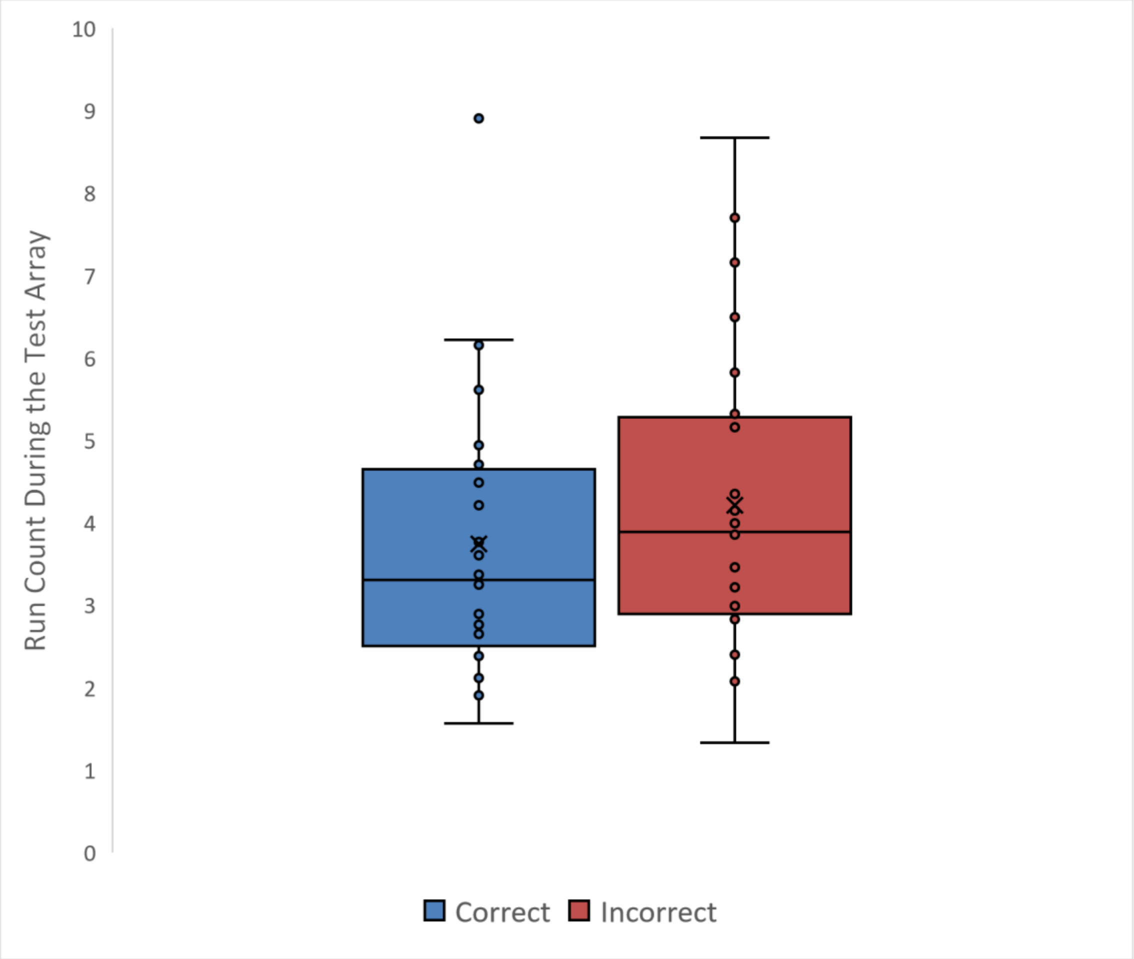


Figure 23: Run Count During the Test Array as a Function of Response Accuracy

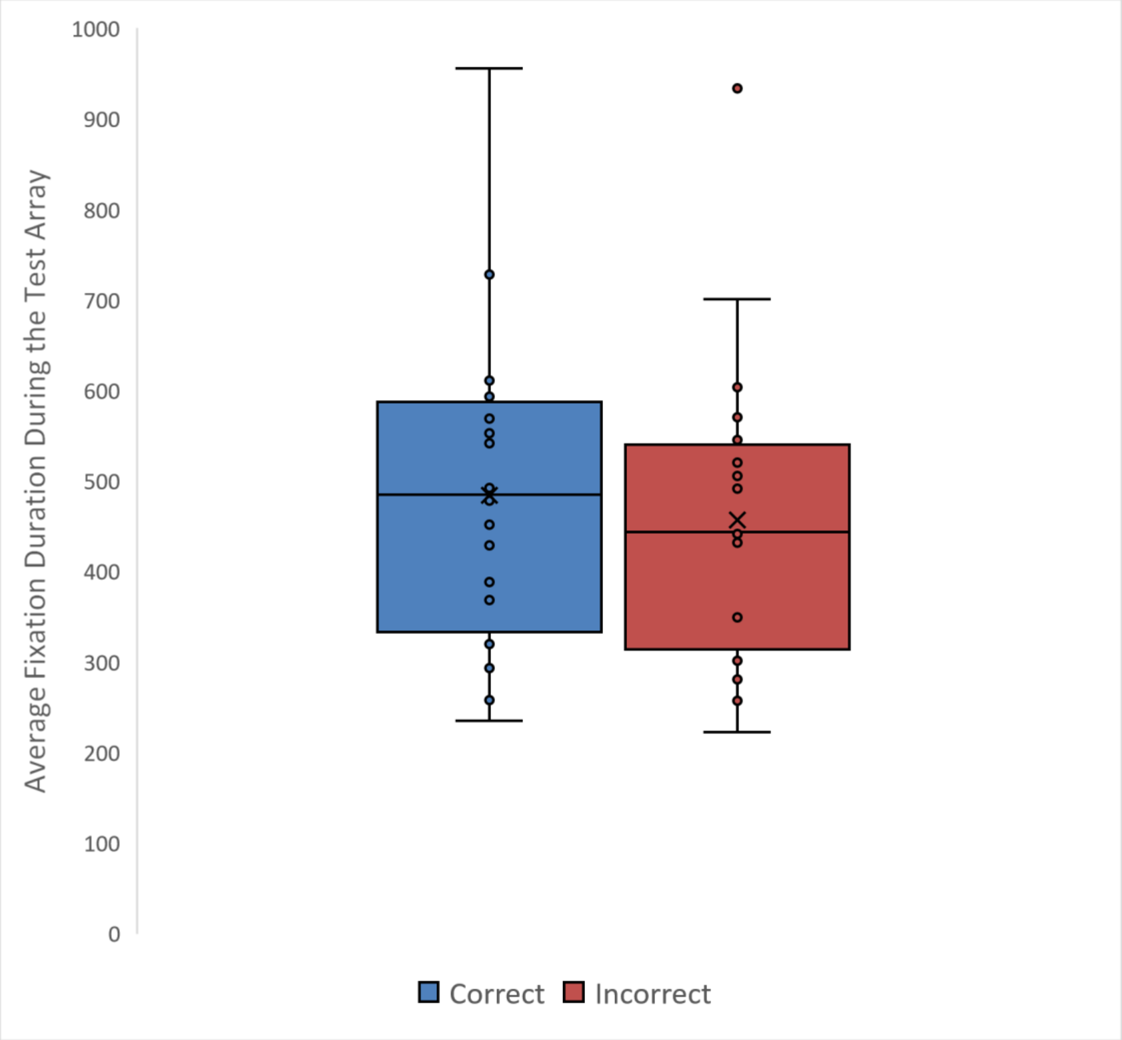


Figure 24: Average Fixation Duration as a Function of Response Accuracy

counts, these results demonstrate that in general, adults made more eye movements and had shorter average fixation for incorrect trials compared to correct trials. Further, the changes in task parameters seemed to have very little if any effect on visual STM performance.

Capacity Estimates

As in Experiment 1, capacity estimates were calculated based on response accuracy at each set size. Pashler's K for set size 6 = 3.56 items. SD = 1.12. Report Pashler's K for set size 9 = 3.4 items. SD = 1.49. Just like in Experiment 1, Pashler's K estimates were averaged to generate a singular K score (3.48) and a bivariate correlation revealed that Pashler's K was significantly negatively correlated with saccade count and run count and positively correlated with average fixation duration (Table 6). Additionally, just like Experiment 1. d' was also calculated for set size 6 (*M*=1.84, *SD*=.868) and for set size 9 (*M*=.989, *SD*=477). Scores were averaged together to generate a single capacity estimate (1.414). Additionally, a bivariate correlation revealed that the averaged d' was significantly negatively correlated with saccade count and run count and positively correlated with average fixation. This is powerful as these results suggest that changing the task, by adopting more infant typical parameters (Ross-sheehy et al., 2003), does not significantly alter capacity estimates in adults, and that the visual dynamics (established in Experiment 1) are still related to a common method of visual STM quantification.

Table 6: Capacity Estimate and Visual Dynamic Correlations (Experiment 3)

Measures	1.	2.	3.	4.	5.
1. Pashler’s K	1	.663**	-.159**	-.074**	.093**
2. dPrime		1	-.157**	-.150**	.132**
3. Run Count			1	.426**	-.307**
4. Saccade Count				1	-.709**
5. Average Fix. Duration					1

** Correlation is significant at the .01 level

Pupil Analyses

Just like Experiment 1, TEPRs were calculated for each individual and they were baseline corrected to minimize any differences in pupil size due to differing luminance values. TEPRs were analyzed using a repeated measures ANOVA with set size (3, 6, 9) and response accuracy (correct, incorrect) as within subjects variables. Results reveal a significant main effect of time, $F(5,115)=2.874$, $p=.018$, $\eta^2_p=.111$, and a marginal effect of accuracy $F(1,23)=3.725$, $p=.066$, $\eta^2_p=.139$. Importantly, these results were qualified by a significant time by response accuracy interaction, $F(5,115)=2.709$, $p=.024$, $\eta^2_p=.105$ (Figure 25). Specifically, TEPR differences between correct and incorrect trials were the greatest early in the test interval and diminished over the duration of the trial period. Follow-up simple effects tests with a Sidak correction for multiple comparisons revealed that there was a significant difference in TEPRs for correct and incorrect trials at 500ms ($p=.004$) and 1500ms ($p=.004$).

Encoding Period Analyses

The encoding period for Experiment 3 was increased from 100ms to 1000ms, (compared to Experiment 1) making it possible to look for evidence that these visual dynamics were predictive of response accuracy during this period. Interestingly, results from a repeated measures ANOVA with change status (change, no-change) and set size (3, 6, 9) as within subjects variables, reveal that nether saccade count ($p=.107$), nor run count ($p=.913$), nor average fixation

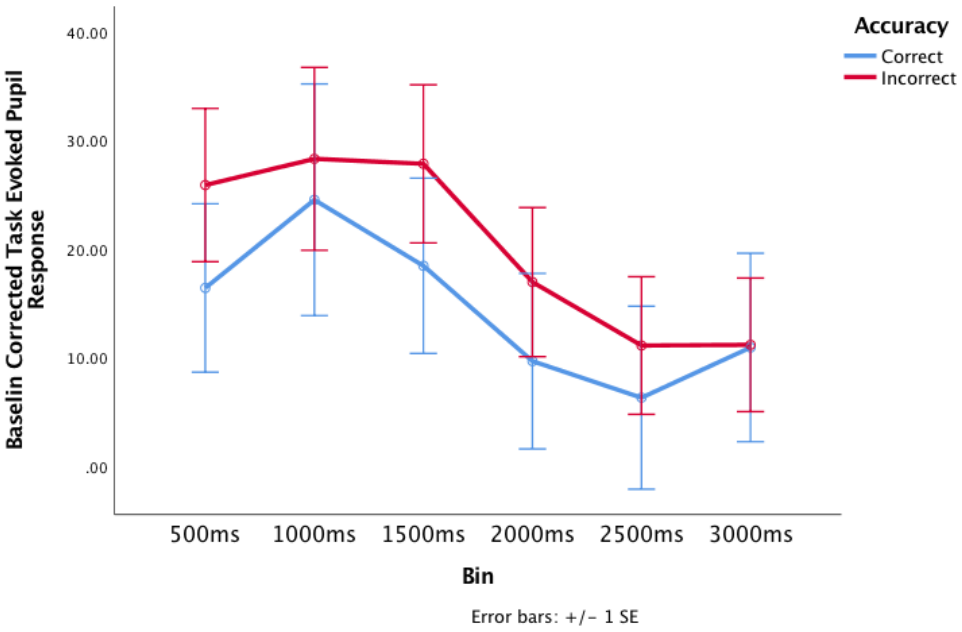


Figure 25: TEPRs as a Function of Response Accuracy and Bin

duration (*p*=.172), nor average pupil size (*p*=.608) during the encoding array were significantly different as a function of response accuracy. This is extremely informative, as it suggests that the significant differences in the identified visual dynamics as a function of response accuracy were not due to on/off task behavior or lapses in attention, or, more importantly, simple individual differences in general behavioral tendencies. Importantly, these results also replicate a similar finding in Ross-Sheehy & Eschman (2019), in that these visual dynamics appear to reflect in-the-moment differences in the quality of STM representations.

Discussion

The goal of Experiment 3 was to test adults in a change-detection task with parameters more typical of those commonly used in infant change-detection tasks—and then to determine if these adaptations are in some way influencing STM performance. Results suggest that testing adults using these infant parameters *did not* alter task performance. Visual dynamics that were identified in Experiment 1 as being predictive of response accuracy, predicted response accuracy here as well. On correct trials, individuals displayed fewer saccades, shorter run counts, and had decreased pupil dilation compared to individuals who responded incorrectly, just like Experiment 1. Increasing the size of the individual items, the total eccentricity, and the encoding time, did not significantly alter task performance. These results provide additional evidence that these visual dynamics reflect visual STM and are not a byproduct of the task parameters. Further, because this task has adopted more infant typical change-detection parameters, it is more likely that these same visual dynamics will be present in infants. Specifically, researchers can feel confident that the similarities between infant visual behavior in Experiment 2A and adult visual behavior in Experiment 1, were not due to the task parameters (i.e. larger items, larger eccentricity, and longer encoding times) used in Experiment 2A.

CHAPTER FIVE

EXPERIMENT 4: A NEW APPROACH TO VISUAL STM ASSESSMENT

Together, Experiments 1-3 have helped identify key visual markers of visual STM in infants and adults and have established the validity of our infant change-detection task as an assessment of visual STM. Experiment 4 aims to further probe the relation between visual behavior during encoding and subsequent change preference by using gaze-contingent eye tracking to control for both attention and recency effects. This also allows us to determine how much change preference scores increase when infants fixate the to-be-changed item during the sample array. In this version of a change-detection task, infants were shown arrays of colored circles just as in the previous experiments. In this version however, all of the arrays included 4 items. Additionally, the location of “to-be-changed” item was pseudorandomly chosen based on fixation order during the sample array. For example, if an infant fixated three items, the change location was equally likely to occur at location 0 (the last item fixated), 1 (the second to last item fixated), or 2 (the third to last item fixated). Infants capable of remembering 1 item should detect the change at *n*-back = 0, whereas infants capable of remembering 2 items should detect the change at both *n*-back = 0 and *n*-back = 1.

This task makes it is possible to know exactly which items were fixated during the encoding array, and perhaps more importantly the order in which they fixated those items. Because of this, visual STM can be quantified in a way that is closely related to adult measures of STM obtained using *n*-back or serial back-presentation tasks. In these tasks, each item in a series is compared to an item that was presented *n* items ago, with *n* denoting serial position relative to test. From this procedure, researchers can estimate capacity based off of the number (*n*) of steps between the initial stimulus and the comparison.

Method

Participants

A total of 62 infants participated in this modified change detection task. Six infants were removed due to fussiness and an additional six were excluded due to lack of interest, one was removed due to an experimenter error, and one was removed due to a technical difficulty. The final sample consisted of 48 infants (22 females and 26 males) that were tested at either 5 or 11 months of age (+- 11 days). All infants were born within three weeks of their due date, and parents reported no birth defects or vision problems. Infant ethnicities were reported as follows: White (n=42), Black (n=1), and two or more races (n=5). Infant names were obtained through the Tennessee Department of Health and Vital Statistics.

Stimuli and Apparatus

The eye tracking set-up was identical to the one described in Experiment 2A. This modified change-detection task was almost identical to the one used in Experiment 2A with two important differences. First, trials included only set size 4 arrays. The location of the color change was determined based on the *n*-back selected for that trial, and could be either 0, 1, or 2. Note: *n*-back for each trial was pseudorandomly chosen based total fixation count during encoding. For example, if an infant only fixated 2 items, *n*-back could be either 0 or 1 (P=.5). If an infant fixated three items, *n*-back could be either 0, 1, or 2 (P=.33). Trials in which infants failed to fixate one array item for at least 200ms were “no change” trials. All other stimuli properties (e.g. size, color, spacing, luminance, etc.) and timing parameters were identical to Experiment 2. Due to the fact that infants failed to consistently fixate more than 2 items, only *n*-back values of 1 and 0, in addition to the no change trials, were included in the analyses.

Design and Procedure

This task incorporated a 2 x 3 design with both condition (change and no change) and *n*-back (no change, 0, and 1) manipulated within subject and age (5- and 11-month-old infants) as a between subject’s variable. As in Experiment 2A, infants were seated on a parent or caregivers lap, approximately 65 cm away from an Eyelink 1000+ eye tracker and were presented a maximum of 80 change-detection trials. However, testing ended sooner if the infant becomes fussy, bored, or fatigued. Also, just like Experiment 2A, primary dependent measures included the proportion of looking to the change circle as a function of total looking (n-back 0 and 1 only), and total looking to the changing versus non-changing arrays as a function of *n*-back. In addition, eye tracking was used to assess pupillometry, saccades, fixations, and run count.

Results

Saccade Count

Saccade count was analyzed via a repeated measures ANOVA with *n*-back (no-change, 0, 1) as a within subjects variables and age (5- and 11-month-old infants) as a between subjects variables. Results are summarized in Table 7. They failed to reveal a significant main effect of *n*-back or age. Although it was hypothesized that saccade count would vary as a function of both age and *n*-back, these results are consistent with Experiment 2A that also failed to produce a significant main effect of change status.

Run Count

Run count was analyzed via a repeated measures ANOVA with *n*-back (no-change, 0, 1) as a within subjects variables and age (5- and 11-month-old infants) as a between subjects variables. Results reveal a significant main effect of age $F(1,43)=5.993$, $p=.019$, $\eta^2_p=.122$, in that the 5-month-old infants had

Table 7: Analysis of Variance Visual Dynamics Results (Experiment 4)

	F	***p***	***df***	***df*$_{error}$**	**η^2**
SACCADE COUNT					
N-Back	0.871	0.422	2	86	0.02
Age	11.682	***0.001***	1	43	0.953
N-Back X Age	0.972	0.382	2	86	0.022
RUN COUNT					
N-Back	81.258	***<.001***	2	86	0.654
Age	5.993	***0.019***	1	43	0.638
N-Back X Age	1.505	0.228	2	86	0.034
AVERAGE FIXATION DURATION					
N-Back	1.73	0.183	2	86	0.039
Age	12.908	***0.001***	1	43	0.231
N-Back X Age	1.18	0.312	2	86	0.027
TEPRs					
Time (Bin)	3.368	***0.006***	5	165	0.093
Age	10.84	***0.002***	1	33	0.247
N-Back	1.801	0.173	2	66	0.052
Bin X Age	0.778	0.567	5	165	0.023
N-Back X Age	1.491	0.233	2	66	0.043
Bin X *N*-Back	1.097	0.364	10	330	0.032
Bin X *N*-Back X Age	0.585	0.689	10	330	0.017

significantly longer runs (*M* = 5.373) than the 11-mos (*M* = 4.403). These results support the notion that increased run count during the test array may be related to a poor/weak memory representation as the 5-month-old infants are producing significantly more saccades during the test array compared to the 11-month-old infants. There was also a significant main effect of *n*-back, $F(2,86)=81.258$, $p<.001$, $\eta^2_p=.654$ (Figure 26). While the age by *n*-back interaction, was not significant, follow-up simple effects tests with a Sidak correction for multiple comparisons showed that both 5- and 11-month-old infants had significantly shorter runs during the test array on the no-change trials (*M* = 3.073) compared to the trial where the changed item was an *n*-back of 0 (*M* = 6.021; $p<.001$) and an *n*-back of 1 (*M* = 6.020; $p<.001$). There was no significant difference in run count for *n*-backs of 0 and 1. These results support the idea that increases in run count may accompany the detection of novelty.

Average Fixation Duration

Average fixation duration was analyzed via a repeated measures ANOVA with *n*-back (no-change, 0, 1) as a within subjects variable sand age (5- and 11-month-old infants) as a between subjects variables. Like saccade count, results failed to reveal a significant main effect of *n*-back or age.

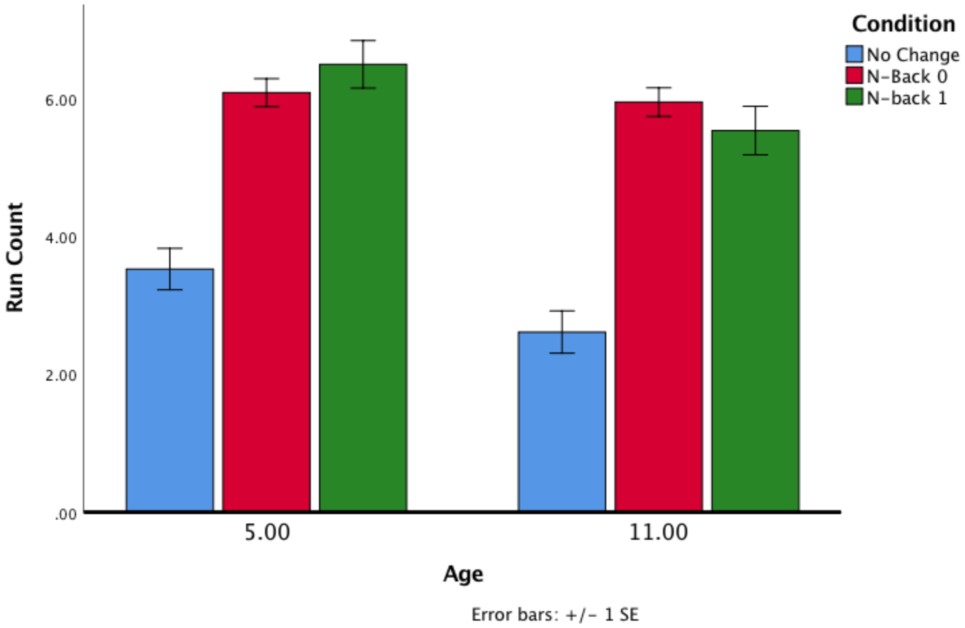


Figure 26: Run Count as a Function of Period, Age, and *N*-Back Value

Change Preference

Change preference was calculated as the proportion of looks to the changed item, as a function of looking to all items. A repeated measures ANOVA, with *n*-back (0, 1) as the within subjects variable and age (5mo and 11mo) as the between subjects variable revealed a significant main effect of n-back $F(1,49)=24.637$, $p<.001$, $\eta^2_p=.335$ with both 5- and 11-month-old infants demonstrating significantly higher change preference scores when the *n*-back was 0, compared to 1 (Figure 27). The age by *n*-back interaction was not significant. To determine if change detection varied as a function of *n*-back, separate one-sample *t*-tests were conducted comparing change preference were to chance (.25) for each age, and *n*-back. Results revealed that 5-month-old infants demonstrated a significant change preference at *n*-backs of both 0 ($t(23)= 9.051$, $p<.001$) and 1 ($t(23)= 3.496$, $p=.002$). Additionally, 11-month-old infants also displayed a significant change preference at both *n*-backs of 0 $t(23)= 13.055$, $p<.001$ and 1 $t(23)= 2.456$, $p=.022$.

In an attempt to assess the relationship between the visual dynamics identified in Experiment 1 and change preference in the current task, bivariate correlations were conducted. Specifically, a bivariate correlation revealed that change preference at an *n*-back of 0 was significantly negatively correlated with saccade count, both during the memory and test array, run count, during the memory array, and IA count during the memory and test array (Table 8).

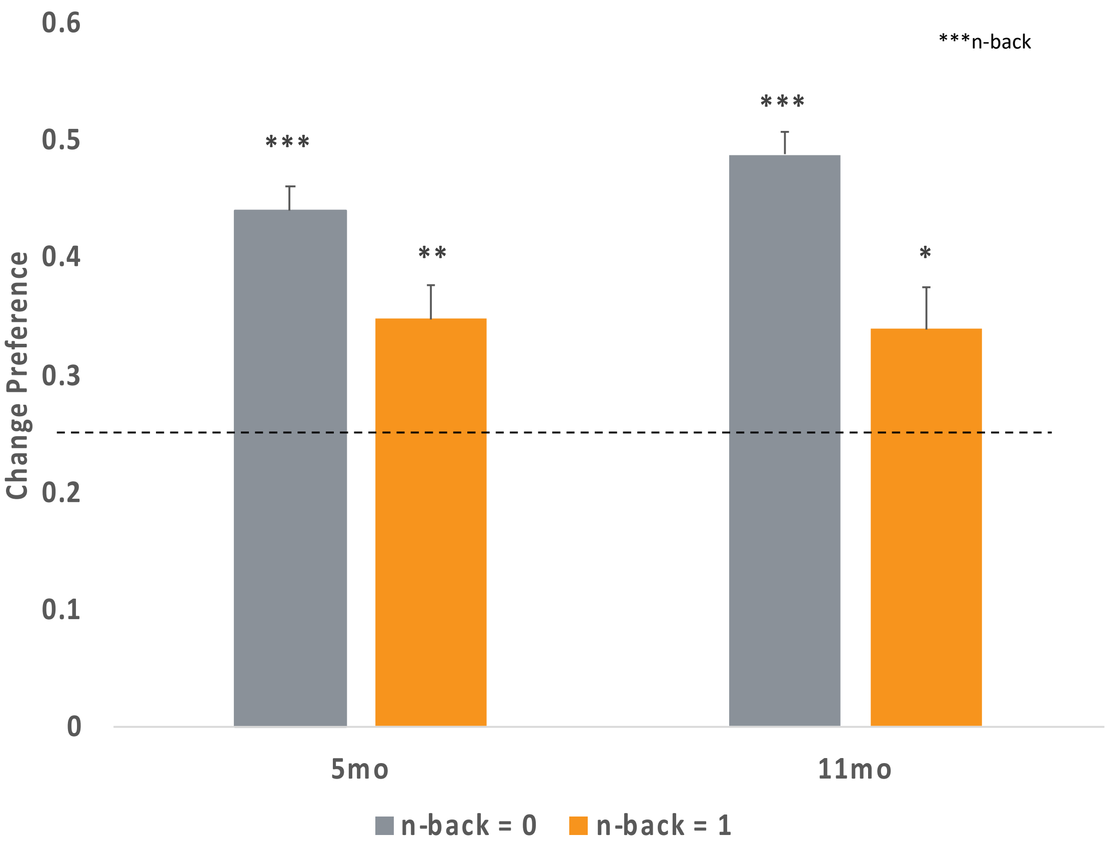
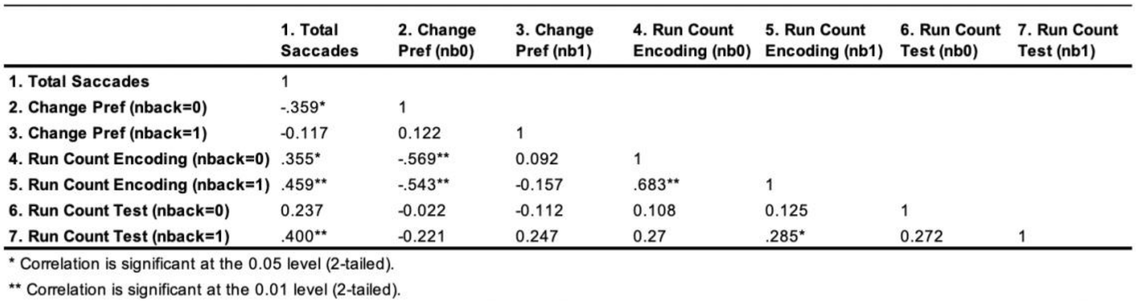


Figure 27: Change Preference and a Function of Age and *N*-Back Value

Table 8: Change Preference and Visual Dynamics (Experiment 4)



	1. Total Saccades	2. Change Pref (nb0)	3. Change Pref (nb1)	4. Run Count Encoding (nb0)	5. Run Count Encoding (nb1)	6. Run Count Test (nb0)	7. Run Count Test (nb1)
1. Total Saccades	1						
2. Change Pref (nback=0)	-.359*	1					
3. Change Pref (nback=1)	-0.117	0.122	1				
4. Run Count Encoding (nback=0)	.355*	-.569**	0.092	1			
5. Run Count Encoding (nback=1)	.459**	-.543**	-0.157	.683**	1		
6. Run Count Test (nback=0)	0.237	-0.022	-0.112	0.108	0.125	1	
7. Run Count Test (nback=1)	.400**	-0.221	0.247	0.27	.285*	0.272	1

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Pupil Analyses

As in all previous experiments, TEPRs were averaged into 500ms bins and baseline corrected (first 100ms of the display array. Just like in Experiment 2A, an outlier analysis revealed 4 outliers. Those participants were removed from the current analysis. A repeated measures ANOVA with bin (500ms, 1000ms, 1500ms) and *n*-back (no change, 0,1) as within subjects variables and age (5mo, 11mos) as the between subjects variable revealed a main effect of bin, *F*(5, 160)=3.451, *p*=.005, η^2_p=.097 (Figure 28). Additionally, these results revealed a significant main effect of age, *F*(1, 32)=9.847, *p*=.004, η^2_p=.235, in that 5-month-old infants had a greater pupil change from baseline (*M* = 92.902) compared to 11-moht-old infants (*M* = 49.211). Finally, these results also revealed a significant main effect of *n*-back, *F*(2, 320)=3.261, *p*=.045, η^2_p=.092, in that pupils during the no-change trials were the largest, followed by trials with an *n*-back of 0, then an *n*-back of 1 (Figure 28). These results are similar those reported in Experiments 1-3 in that pupils consistently show the largest change form baseline at the smallest set sizes/ *n*-back. Further, these results suggest that pupils again are a useful tool in identifying meaningful differences in task performance as a result of trial differences.

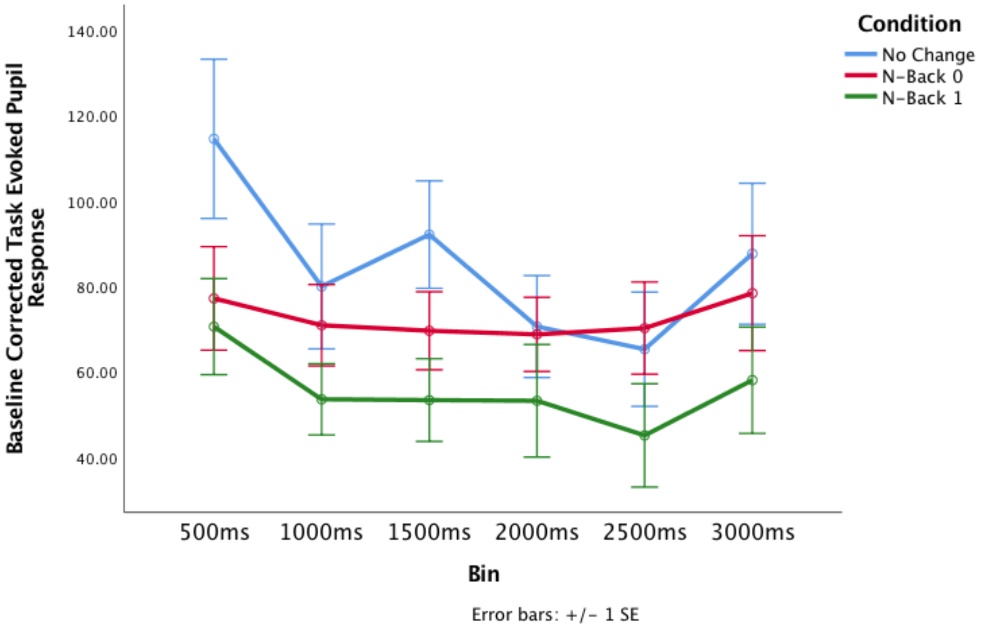


Figure 28: TEPRs as a Function of Age and *N*-Back Value

Discussion

Ross-Sheehy and Eschman (2019), demonstrated that having fixated the to-be-changed item significantly increases change preference. Results from Experiment 4 replicate Ross-Sheehy and Eschman (2019) and demonstrate fixation during the sample predicts change preferences scores. Results from this experiment further extend this work and demonstrate that the sequential order of fixations during the sample array can influence a preference for the changed item during the test array, with significantly stronger change preference scores for the last item fixated during encoding. Thus, just like Ross-Sheehy and Eschman (2019), these results further suggest that change preference scores are strongly influenced by incidental attention to the to-be-changed item during the sample array. Although this finding suggests change preference may not be an ideal score for assessing memory for the entire array, the presence of these serial order effects is clear evidence that STM for the memory array influences attention during the test array. Additionally, similar to Experiment 2, run count was significantly influenced by task demands. Infants had significantly longer runs on change trials than they did on the no-change trials.

CHAPTER SIX
GENERAL DISCUSSION AND CONCLUSIONS

The results of this project have significantly impacted our general understanding of the development of infant visual STM. All ANOVA results are summarized in Table 9. Experiment 1 revealed quantifiable looking dynamics that are easily measured and appear to predict visual STM accuracy in adults. These measures include saccade count, run count, average fixation duration, and TEPRs. Prior to this work, adult visual STM has typically been quantified based solely on accuracy measures derived from overt behavioral response (i.e. button and/or verbal response). Although accuracy is an effective tool for identifying individual differences in things like STM capacity, these types of behavioral responses are limiting, in that they require a dichotomous yes/no response. New measures, such as those tested here, provide a more continuous metric of visual STM, allowing for a more *fine-grained* analysis of the quality of memory representations. In addition to providing a more complete picture regarding the contents of memory, this way of measuring visual STM may also make it possible to 1) identify individual differences in visual STM performance in a population that is incapable of providing an overt behavioral response (infants), and 2) allow for an examination of underlying similarities and differences between infant and adult mechanisms of change detection. Importantly, these results suggest that low-level visual dynamics can reveal measurable “markers” of change detection in adult participants, perhaps even before they make a behavioral decision.

Table 9: Summary Table of All Significant Differences in Visual Dynamics

*p<.05, **p<.01, ***p<.001, † = marginal

	Exp. 1	Exp. 2A	Exp. 2B	Exp. 3	Exp. 4
SACCADE COUNT					
Set Size/N-Back	**	***	***		
Age		***			**
Change Status					
Accuracy	*			*	
Set Size X Change Status					
Change Status X Age					
Set Size X Age					
Set Size X Accuracy					
Change Status X Accuracy					
Set Size X Change Status X Accuracy					
Set Size X Change Status X Age					
RUN COUNT					
Set Size/ N-Back	***	***	***	***	***
Age					*
Change Status		***	***		
Accuracy				*	
Set Size X Change Status			*		
Change Status X Age					
Set Size X Age					
Set Size X Accuracy					
Change Status X Accuracy					
Set Size X Change Status X Accuracy					
Set Size X Change Status X Age					
AVERAGE FIXATION DURATION					
Set Size/ N-Back		***	***		
Age		**			**
Change Status					
Accuracy	*			†	
Set Size X Change Status			*		
Change Status X Age					
Set Size X Age					
Set Size X Accuracy					
Change Status X Accuracy					
Set Size X Change Status X Accuracy					
Set Size X Change Status X Age					
TEPRs					
Time (bins)	***		***	*	**
Age					**
Accuracy	*			†	
Set Size/ N-Back			*		
Change Status		***			
Bin X Set Size		*	***		
Bin X Accuracy				*	
Bin X Change Status		*			
Set Size X Change Status					
Bin X Set Size X Change Status			**		

Experiment 2A has demonstrated that these memory-dependent visual dynamics are detectable in infant populations as well. Specifically, saccade counts, run counts, mean fixation duration and TEPRs during the test array differed systematically as a function of condition and set size. Although it is impossible to know if these dynamics reflect active STM maintenance, there are several reasons to suspect they do. First, much like the adults in Experiment 1, visual dynamics differed as a result of changes in set size, suggesting that the contents of a visual scene (and perhaps the contents of memory) can change the way infants view that scene. This is important because we know that adults display certain visual dynamics when they do not have a solid memory representation which suggests a lack of visual STM maintenance, and infants display a similar trend. Second, and in support of this claim, run count varied as a function of change status. Specifically, infants made significantly longer runs at all set sizes, on trials that contained a change. Again, this suggests that infant visual dynamics are influenced by the contents of a visual scene. Finally, when adults were run in the same passive change-detection task (Experiment 2B), they displayed virtually identical visual behaviors as the infants. The most important of these effects is the fact that adults, just like infants, displayed significantly longer runs on change trials. This is important as the results from Experiment 1 demonstrated that several of these visual behaviors did indeed vary with accuracy. Experiments 1 and 2 have identified a meaningful relationship between visual dynamics and visual STM performance in a change-detection task. Due to the fact that there were considerable differences in the properties of the tasks used in Experiments 1 and 2, it was possible that the bigger items, larger total eccentricities, and longer encoding time artificially inflated performance. As a result, it was important to demonstrate that these same effects could be found in an experiment that adopted several of these same parameters and also included a button response to measure accuracy.

Experiment 3 successfully replicated the main results of Experiment 1, in spite of significant changes to the parameters of the change-detection task. Specifically, Experiment 3 (the hybrid task) adopted a more *infant typical* approach to measuring visual STM in adults, with the addition of a behavioral response. Just like Experiment 1, adult visual dynamics varied as a function of both set size, change status, and response accuracy. Specifically, adults displayed longer runs, more saccades, and larger TEPRs on trial where they responded incorrectly. Thus far, these 3 experiments have all provided evidence that visual dynamics change as a result of a set size and change status. Specifically, increased saccade and run counts appear to coincide with relatively poor STM representations, suggesting increased visual scanning during the test array may be necessary in order to make a correct response.

This task also includes set sizes that are commonly outside of typical adult capacity estimates (Luck & Vogel, 1997). So, in addition to adopting more infant typical parameters (like Experiment 2), this task also includes larger set sizes (like Experiment 1). This accomplishes the important goal of increasing the power of the incorrect trials. Ross-Sheehy and Eschman (2019) report that response accuracy in their task was high (nearly 96% correct). Experiment 3 reduces that number to 77%, thus providing a significant increase in the number of incorrect responses.

Additionally, these results further clarify the relation between visual dynamics and STM accuracy by demonstrating that low-level visual dynamics during the *encoding* portion of a change-detection task do not influence subsequent visual STM performance. If visual dynamics during the test array were merely driven by individual differences in things like mean fixation duration (Colombo, Mitchell, Coldren, & Freeseman, 1991), then we would expect this relationship to hold during the encoding array as well. Importantly, it does not. Visual dynamics like saccade count, run count, and TEPRs appear to vary during the test array, presumably while the subject is comparing the contents of WM to the current array. Also, these results suggest that increasing the encoding duration to mimic typical infant parameters does not influence visual STM performance—neither does increasing the size of the individual items nor the total eccentricity of the array.

Experiment 4 examined the extent to which change-detection in infancy is influenced by serial fixation effects during the sample array. Results revealed clear sequential fixation effects, with strongest change preference scores for the most recently fixated item. Moreover, change preference scores were significantly higher for the last item fixated (*n*-back = 0) then for the second-to-the-last item fixated (*n*-back = 1). Importantly, just like Experiment 2, run count was significantly shorter for trials that provided a smaller cognitive load (i.e. no-change trials) and scaled as the difficulty increased (*n*-back of 0 then 1 respectively). The TEPRs in Experiment 4 add evidence to support this claim as pupil size also changed as a result of change status. Pupils changed the most from baseline on the trials that did not change (perhaps a form of novelty detection) and were the largest for trails containing and *n*-back of 1. These results are similar to findings in Experiment 2 that pupils were the largest at the smallest set sizes and further posit that TEPRs are a meaningful metric when it comes to measuring visual STM performance. In addition to providing clear evidence of meaningful differences in visual dynamics as a result of trial differences and recency effects, these results demonstrate the efficacy of using an *n*-back approach as a means of assessing visual STM capacity.

Is Change Preference Appropriate for Estimating Visual STM Capacity?

At the outset of this project, there was hope that by combining visual dynamics with change preference scores, it could be possible to compute visual STM capacity in infants, much the same way it is calculated in adults. Recently however, Ross-Sheehy and Eschman (2019), demonstrated change preference was driven predominantly by attention during the sample array, with strong change preference on trials in which infants happened to fixate the to-be-changed item during the sample array. Building off of this, the current set of experiments successfully replicated these findings, and further posited that the sequential order of fixations during the sample array *does* significantly influence preference for the changed item during the test array in a change detection task. Results from this dissertation support the notion that change preference, as calculated in these single-screen, one-shot tasks, is not a reliable predictor of visual STM ability in either infants or adults. Given the relation between change preference and fixation during the sample array, it is reasonable to assume that infants who are fast scanners are more likely to have fixated the to-be-changed item, and thus more likely to demonstrate a significant change preference. However, as in Ross-Sheehy and Eschman (2019), results from Experiment 3 cast doubt on this hypothesis, as saccade and run counts during the encoding array were not related to change preferences at test. Thus, although infants need not encode the entire array in order to demonstrate a change preference, fixation behavior during encoding can substantially alter the magnitude of change preference.

Can You Use an *N*-Back Approach to Estimate Visual STM Capacity?

In an attempt to more specifically quantify serial order effects on subsequent change preference scores, Experiment 4 systematically randomized the location of the changed item based on fixation order during the sample array. Results demonstrate the utility of this type of visual STM measure. Specifically, by controlling the item that changes, researchers can directly probe capacity. Further, they can test interference effects by examining the strength of the memory representation, with recently fixated items resulting in significantly greater change preference scores.

The current task may also provide answers to the question of whether participants are employing some type of ensemble or configural processing in one-shot, change-detection tasks. Because the individual items are isolated, and researchers can control how long ago they viewed the to-be-changed item, it is likely that differences in visual STM performance are due to the ability to recall individual components of the array. At worst, this suggests that the individual components that make up a configural image have a stronger memory representation. At best, this may suggest that participants are encoding the individual items within a scene and not participating in any type of configural processing. Before researchers can make this claim, however, more work needs to be done, but this approach could be our best bet at resolving this issue.

Finally, the current *n*-back approach offers a method of assessing visual STM performance in infants that directly parallels a common method of adult STM assessment. The *n*-back approach is a task that has been difficult to use in infant populations, but the research in this dissertation suggests that it is possible, and importantly, researchers can gain meaningful information with regard to visual STM performance. This in itself is a significant contribution and could change the way that we measure visual STM or WM from a developmental standpoint—offering the first means of assessing visual STM capacity from infancy through adulthood.

Are Passive Change-Detection Tasks Fundamentally Different than those that Have an Overt Behavioral Response?

This series of experiments provides a significant amount of evidence supporting the idea that adult measures of visual STM or visual WM *are* similar to infant *passive* measures of visual STM—as measured via one-shot change-detection tasks. Both infants and adults display several visual similarities that change as a function of set size and task difficulty (e.g. saccade count, run count, average fixation duration, and TEPRs). Further, adult visual dynamics in a passive change-detection task were nearly identical to an active change-detection task involving an explicit decision (“same” or “different”) along with an overt behavioral response (button press). Future work will be needed to address the idea that adults could have been explicitly noticing “change” and “no change” even during the passive task. Currently however, these data have provided strong evidence that 1) adding a button press does not measurably alter the task, and 2) adult and infant measures of visual STM are *qualitatively,* and to some extent, *quantitatively* similar.

Using Visual Dynamics in the Future

The visual dynamics specified in this dissertation (saccade and run counts, mean fixation duration, and TEPRs), are easily measured regardless of age, using standard eye tracking procedures. These dynamics were identified in a direct replication of a well-known measure of visual STM capacity and importantly, were predictive of response accuracy. Experiments 2-4 aimed to identify these same dynamics in a variety of change-detection tasks at multiple ages. The results suggest that not only are they present in both infant and adult participants, but they are affected by the task demands in a very similar way. These dynamics are telling us something very specific with regard to how visual STM affects the visual behavior in both infants and adults. Specifically, when the memory representations are weak, adults and likely infants demonstrate more saccades, they have longer runs, shorter average fixations, and their pupils change more from baseline. This suggests that visual dynamics such as those highlighted here reflect similar implicit visual STM mechanisms across both infant and adult populations, as well as across implicit and explicit change-detection tasks.

In addition to providing evidence of the underlying processes(es) involved in successful visual STM performance, the visual dynamics that have been explored in this series of experiments demonstrate have proven to be effective in demonstrating both group and individual differences in visual STM performance in a change-detection task. Specifically, this dissertation has demonstrated that similar individual differences exist in both infant and adult populations. The dynamics outlined in this dissertation can easily be viewed on an individual level, providing robust differences from one participant to the next. Future work should also be aimed at longitudinal outcomes that may be predicted but individual differences in visual dynamics early in development. This would allow researchers to develop targeted intervention techniques that may be applied during a time where they are most likely to be successful.

Due to the success of identifying both group and individual differences in these visual dynamics, future work should be aimed at applying these dynamics to other tasks that measure foundational cognitive abilities. Specifically, in tasks where visual STM may mediate performance, these visual dynamics may shed additional light onto task performance and perhaps more importantly, individual differences. Additionally, these results demonstrate the utility of taking a more fin-grained approach to exploring looking behavior in both adults and infants. As eye tracking technologies continues to improve, researchers have an incredible opportunity to look at their data in ways not possible previously. This fine-grained approach to offers more individual differences increasing the likelihood of identifying meaningful predictors of later cognition early in development.

General Conclusions

This project has provided several critical insights significantly advancing our understanding of how infants learn about the world around them as well as how these early learning strategies may develop into critical thinking skills throughout childhood. Specifically, the low-level visual dynamics that have been highlighted in this dissertation seem to be consistently reflecting the quality of the memory representation for both infant and adult participants in a one-shot, change-detection task. The results that have been discussed in this series of studies provide strong evidence that infant and adult measures of visual STM may be accessing a similar memory mechanism. Both infants and adults displayed similar visual dynamics across multiple versions of a change-detection task. Importantly, these dynamics were consistent with those displayed in tasks with an objective measure of accuracy. These results provide significant progress toward a unified theory of visual STM development across the lifespan.

LIST OF REFERENCES

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). Published by : Wiley on behalf of the Society for Research in Child Development The Cognitive and Behavioral Characteristics of Children With Low Working Memory Tracy Packiam Alloway Susan Elizabeth Gathercole and University of Stirling Hannah Kirkwood. *Child Development*, *80*(2), 606–621.

Alloway, T. P., & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences*, *21*(1), 133–137. https://doi.org/10.1016/j.lindif.2010.09.013

Alvarez, G., & Cavanagh, P. (2010). The capacity of visual short-term memory is set by total information load, not number of objects. *Journal of Vision*, *2*(7), 273–273. https://doi.org/10.1167/2.7.273

Awh, E., Barton, B., & Vogel, E. K. (2014). *Visual Working Memory a Fixed Number of Represents Items Regardless of Complexity*. *18*(7), 622–628.

Baddeley, A. ., & Logie, R. H. (1999). Mechanisms of active maintenance and executive control. In A. Miyake & P. Shah (Eds.), *Working memory: A multiple-component model. Models of working memory* (pp. 28–61).

Baddeley, A. D., & Hitch, G. J. (1974). Working Memory. In *Psychology of learning and motivation* (pp. 47–89).

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*(4), 485.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276–292. https://doi.org/10.1037//0033-2909.91.2.276

Bettencourt, K. C., & Xu, Y. (2016). Decoding the content of visual short-term memory under distraction on occipital and parietal areas. *Nature Neuroscience*, *19*(1), 150. https://doi.org/10.1016/S2215-0366(16)30284-X.Epidemiology

Blaga, O. M., & Colombo, J. (2006). Visual processing and infant ocular Latencies in the overlap paradigm. *Developmental Psychology*, *42*(6), 1069–1076. https://doi.org/10.1037/0012-1649.42.6.1069

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*(3), 205–228. https://doi.org/10.1080/87565640801982312

Colombo, J., Mitchell, D. W., Coldren, J. T., & Freeseman, L. J. (1991). Individual Differences in Infant Visual Attention: Are Short Lookers Faster Processors or Feature Processors ? Author (s): John Colombo , D . Wayne Mitchell , Jeffrey T . Coldren and Laura J . Freeseman Published by : Wiley on behalf of the Society for. *Child Development*, *62*(6), 1247–1257.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), S0140525X01003922. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2005). *Working Memory Capacity* (1st ed.). New York, NY: Psychology Press.

Daneman, M., & Carpenter, P. A. (2004). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*. https://doi.org/10.1016/s0022-5371(80)90312-6

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*. https://doi.org/10.1111/1467-8721.00160

Fagan, J. F. (1984). The intelligent infant: Theoretical implications. *Intelligence*, *8*(1), 1–9. https://doi.org/10.1016/0160-2896(84)90002-3

Feigenson, L. (2007). Continuity of format and computation in short-term memory development. In *Short- and long-term memory in infancy: Taking the first steps toward remembering*. Retrieved from papers3://publication/uuid/FBA843B3-34F6-44DC-8997-B2C8AFB3D144

Goldinger, S. D., & Papesh, M. H. (2012). Pupil Dilation Reflects the Creation and Retrieval of Memories. *Current Directions in Psychological Science*, *21*(2), 90–95. https://doi.org/10.1177/0963721412436811

Goldwater, B. C. (1972). Psychological significance of pupillary movements. *Psychological Bulletin*, Vol. 77, pp. 340–355. https://doi.org/10.1037/h0032456

Gredebäck, G., Eriksson, M., Schmitow, C., Laeng, B., & Stenberg, G. (2012). Individual differences in face processing: Infants’ scanning patterns and pupil dilations are influenced by the distribution of parental leave. *Infancy*, *17*(1), 79–101. https://doi.org/10.1111/j.1532-7078.2011.00091.x

Heaver, B., & Hutton, S. B. (2011). Keeping an eye on the truth? pupil size changes associated with recognition memory. *Memory*, *19*(4), 398–405. https://doi.org/10.1080/09658211.2011.575788

Hellmer, K., Soderlund, H., & Gredeback, G. (2016). The eye of the retriever: Developing episodic memory mechanisms in preverbal infants assessed through pupil dilation. *Developmental Science*, (September), 1–10. https://doi.org/10.1111/desc.12520

Hess, E. H. (1975). *The tell-tale eye:How your eyes reveal hidden thoughts and emotions*. Von Nostrand Reinhold.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem solving. *Science*, *143*(3611), 1190–1192.

Johnson, M. H. (2008). The Development of Visual Attention: A Cognitive Neuroscience Perspective. In *Brain Dovelopement and Cognition*. https://doi.org/10.1002/9780470753507.ch9

Kwon, M.-K., Luck, S. J., & Oakes, L. M. (2013). Visual short-term memory for complex objects in 6-and 8-month-old infants. *Child Development*, *70*(4), 646–656. https://doi.org/10.1002/ana.22528.Toll-like

Luck, S. J. (2007). Visual short-term memory. In *Scholarpedia* (2nd ed., p. 3328).

Luck, S. J. (2008). Visual short-term memory. In S. J. Luck & A. Hollingworth (Eds.), *VIsual Memory* (pp. 43–85).

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

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. https://doi.org/10.1038/nn.3655

Oakes, L., Ross-Sheehy, S., & Luck, S. (2007). The development of visual short-term memoryin infancy. In L. Oakes & P. J. Bauer (Eds.), *Short- and long-term memory in infancy and early childhood: Taking the first steps toward remembering.* New York: Oxford University Press.

Oakes, L.M., Ross-Sheehy, S., & Luck, S. J. (2006). Rapid development of feature binding in visual short-term memory. *Psychological Science*, *17*(9), 781–787. https://doi.org/10.1111/j.1467-9280.2006.01782.x

Oakes, Lisa M., Baumgartner, H. A., Barrett, F. S., Messenger, I. M., & Luck, S. J. (2013). Developmental changes in visual short-term memory in infancy: Evidence from eye-tracking. *Frontiers in Psychology*, *4*(OCT), 1–13. https://doi.org/10.3389/fpsyg.2013.00697

Oakes, Lisa M., Hurley, K. B., Ross-Sheehy, S., & Luck, S. J. (2011). Developmental changes in infants' visual short-term memory for location. *Cognition*, *118*(3), 293–305. https://doi.org/10.1016/j.cognition.2010.11.007.Developmental

Oakes, Lisa M., Messenger, I. M., Ross-Sheehy, S., & Luck, S. J. (2009). New evidence for rapid development of colour-location binding in infants' visual short-term memory. *Visual Cognition*, *17*(1–2), 67–82. https://doi.org/10.1080/13506280802151480

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*(4), 369–378. https://doi.org/10.3758/BF03210419

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*(2), 283–290. https://doi.org/10.3758/BF03203943

Posner, M. I., & Dehaene, S. (1994). Attentional networks. *Trends in Neurosciences*. https://doi.org/10.1016/0166-2236(94)90078-7

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2001). Visual short-term memory in the first year of life: capacity and recency effects. *Developmental Psychology*, *37*(4), 539–549. https://doi.org/10.1037/0012-1649.37.4.539

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2002). Processing speed in the 1st year of life: a longitudinal study of preterm and full-term infants. *Dev Psychol*, *38*(6), 895–902. https://doi.org/10.1037/0012-1649.38.6.895

Ross-Sheehy, S., & Eschman, B. (2019). Assessing visual STM in infants and adults: eye movements and pupil dynamics reflect memory maintenance. *Visual Cognition*, *0*(0), 1–15. https://doi.org/10.1080/13506285.2019.1600089

Ross-Sheehy, S., & Newman, R. S. (2015). Infant auditory short-term memory for non-linguistic sounds. *Journal of Experimental Child Psychology*, *132*, 51–64. https://doi.org/10.1016/j.jecp.2014.12.001

Ross-sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The Development of Visual Short-Term Memory Capacity in Infants. *Child Development*, *74*(6), 1807–1822.

Ross-sheehy, S., Oakes, L. M., & Luck, S. J. (2011). Exogenous attention influences visual short-term memory in infants. *Developmental Science*, *14*(3), 490–501. https://doi.org/10.1111/j.1467-7687.2010.00992.x.Exogenous

Shaffer, W., & Shiffrin, R. M. (1972). Rehearsal and storage of visual information. *Journal of Experimental Psychology*, *92*(2), 292–296. https://doi.org/10.1037/h0032076

Simmering, V. R. (2011). The development of visual working memory capacity during early childhood. *Journal of Experimental Child Psychology*, *111*(4), 695–707. https://doi.org/10.1016/j.jecp.2011.10.007

Simmering, V. R. (2016). I. Working Memory Capacity in Context: Modeling Dynamic Processes of Behavior, Memory, and Development. *Monographs of the Society for Research in Child Development*, *81*(3), 7–24. https://doi.org/10.1111/mono.12249

Simmering, V. R., & Perone, S. (2013). Working Memory Capacity as a Dynamic Process. *Frontiers in Psychology*, *3*(January), 1–26. https://doi.org/10.3389/fpsyg.2012.00567

Sirois, S., & Jackson, I. R. (2012). Pupil dilation and object permanence in infants. *Infancy*, *17*(1), 61–78. https://doi.org/10.1111/j.1532-7078.2011.00096.x

Stampe, D. M. (1993). Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. *Behavior Research Methods, Instruments, & Computers*, *25*(2), 137–142. https://doi.org/10.3758/BF03204486

Unsworth, N., & Engle, R. W. (2007). On the Division of Short-Term and Working Memory: An Examination of Simple and Complex Span and Their Relation to Higher Order Abilities. *Psychological Bulletin*, *133*(6), 1038–1066. https://doi.org/10.1037/0033-2909.133.6.1038

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic Bulletin & Review*, *22*(3), 757–765. https://doi.org/10.3758/s13423-014-0747-6

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(4), 601–615. https://doi.org/10.3758/s13415-016-0417-4

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503. https://doi.org/10.1038/nature04171

Wetzel, N., Buttelmann, D., Schieler, A., & Widmann, A. (2016). Infant and adult pupil dilation in response to unexpected sounds. *Developmental Psychobiology*, *58*(3), 382–392. https://doi.org/10.1002/dev.21377

Wickens, T. (2001). *Elementary signal detection theory*. Oxford University Press. New York.

APPENDIX

Table 10: Experiment 1-4 RGB and Luminance Values

		Color 1	Color 2	Color 3	Color 4	Color 5	Color 6	Color 7	Color 8	Color 9	Color 10	Background Color
Experiment 1	*Color*	Black	Blue	Green	Red	Purple	White	Yellow	---	---	---	Grey
	RGB	0, 0, 0	46, 62, 124	24, 119, 24	149, 27, 27	119,33,119	255,255,255	181,181,14	---	---	---	180, 180, 180
	Luminance	123.9	142.4	154	138.5	142	215.4	197.5	---	---	---	173.6
Experiment 2	*Color*	Red	White	Blue	Yellow	Green	Purple	Orange	Teal	Brown	---	Grey
	RGB	241, 0, 0	255,255,255	0, 0, 255	255, 248, 0	0, 153, 51	129, 0, 130	255, 120, 0	0, 221, 225	129, 85, 40	---	116, 116, 116
	Luminance	46.27	263.8	26.55	238.8	94.88	29.82	104.7	185.8	58.19	---	86.03
Experiment 3	*Color*	Red	Orange	Yellow	Green	Teal	Blue	Light Purple	Pink	Maroon	Plum	Grey
	RGB	255, 0, 0	255, 153, 51	255, 255, 0	0, 53, 0	0, 255, 255	0, 0, 255	153,51,255	255,51,255	102, 0, 51	51, 0, 102	180, 180, 180
	Luminance	54.67	149.2	243	91.89	184.9	23.68	69.37	93.16	20.4	13.19	173.6
Experiment 4	*Color*	Red	White	Blue	Yellow	Green	Purple	Orange	Teal	Brown	---	Grey
	RGB	241, 0, 0	255,255,2555	0, 0, 255	255, 248, 0	0, 153, 51	129, 0, 130	255, 120, 0	0, 221, 225	129, 85, 40	---	116, 116, 116
	Luminance	46.27	263.8	26.55	238.8	94.88	29.82	104.7	185.8	58.19	---	86.03

VITA

Bret Eschman was born in Belleville, Illinois, to the parents of Greg and Vicki Eschman. He has a one sibling, a younger sister, Bethany. He attended High Mount Elementary School and then continued to Belleville West High School in Belleville, Illinois. After graduation, he headed north to Illinois College where he was introduced to experimental psychology. During his time at Illinois College, Bret worked in a research lab under the direction of Dr. Jeremy Turner. He obtained a Bachelor of Science degree in psychology and sociology from Illinois College in 2011. After graduation, he accepted a teaching assistantship at Missouri State University in the psychology department. During his time at Missouri State, he worked with Dr. D. Wayne Mitchell in the Infant Learning and Perception Lab as a research assistant. Bret graduated with a Master of Science degree in Experimental Psychology in 2013. After receiving his master's degree, Bret accepted a research assistantship at East Tennessee State University in Johnson City, Tennessee. After two years in the program, his major advisor accepted a position at the University of Tennessee, Knoxville and Bret made the move with her. During the past five years, Bret has been a research Assistant to Dr. Shannon Ross-Sheehy in the Infant Visual Cognition Lab. He plans to graduate from the University of Tennessee with a Doctor of Philosophy degree in Cognitive and Developmental Psychology in August of 2019.