EXAMINING FACTORS ASSOCIATED WITH VISUOSPATIAL WORKING MEMORY DEVELOPMENT FROM INFANCY THROUGH THE PRESCHOOL YEARS

Laura Jean Zimmermann, M.A.

Thesis Advisor: Rachel F. Barr, Ph.D.

ABSTRACT

In early childhood, the development of executive functioning (EF) takes place. This complex system comprised of working memory, inhibitory control, and shifting is crucial for successful cognitive and socio-emotional development. A subcomponent of working memory is visuospatial working memory (VSWM). VSWM is defined as maintaining, updating, and manipulating visual and spatial information in the short-term, while ignoring irrelevant yet potentially salient distracting information. This ability emerges early in infancy and develops throughout the preschool years.

This dissertation examines the development of VSWM of children between 18-months and 5-years of age on several tasks with varying levels of complexity. Both established object-based and newly developed touchscreen tasks are considered. Chapter two examines factors associated with performance on a simple 3-location object occlusion VSWM task, called *Hide the Pots*, by 18- and 24-month-olds. In study 1, infants performed significantly worse after a 10s delay than after a 2s delay. Performance was associated with individual differences in productive vocabulary. In study 2 removing color as a cue did not impact performance by 18-month-olds.

Chapter 3 examines factors associated with performance on a complex complex 8-location object occlusion VSWM task called *Spin the Pots* in toddlerhood through preschool. There were age-related changes between 2- and 4-years of age; 4-year-olds performed significantly better on the task with girls out-performing boys. Chapter four developed a new task, *Find the Stars*, based upon the CANTAB Spatial Working Memory task with 3- to 5-years to provide a nonverbal and
age-appropriate touchscreen-delivered measure of VSWM. The extent to which age and cognitive load effect performance on the working memory task and its’ relationship to inhibitory control, motor-spatial imitation, and parent report EF measures were assessed. There were age-related changes, associations on all three tasks, and a relationship between performance and parent report EF measures.

The present dissertation demonstrates multiple factors associated with developmental changes in VSWM in the first five years of life. They include delay, vocabulary, visual perceptual cues, cognitive load, and other EF components. These findings have important implications for the assessment of VSWM and factors that relate to the trajectory of VSWM in early childhood.
ACKNOWLEDGEMENTS

This dissertation would not have been possible without the generous support of so many people. First, I would like to thank my committee, Dr. Chandan Vaidya, Dr. Andrei Medvedev, and Dr. Daniel Hyde for their assistance and encouragement throughout this dissertation process. Your insightful comments and helpful critiques were extremely valuable to me.

Thank you to the internal and external dissertation funding sources including the Society for Research in Child Development Dissertation Funding Award and the Pilot Research Grant from Georgetown University.

I would like to thank everyone, past and present, in the Early Learning Project, including Emily Oot, Amanda Grenell, Dr. Natalie Brito, and Sylvia Rusnak, for their help, support, and encouragement. I would also like to thank the many research assistants that I have worked with over the years including Hannah Frank, Chagai Mendelson, Sydney Simmons, Christina Sharkey, Herietta Lee, Lawrie Green, Tish Kashani, and Maria Rivera. A special thank you to Ruthie Speidel, Courtney Nugent, and Julia Tonnessen for their dedication to and enthusiasm for this project.

I would like to thank my former mentors Dr. Richard Aslin, Dr. Celeste Kidd, Dr. Susan Carey, and Dr. Nathan Winkler-Rhoades for preparing me for my PhD training and for the advice along the way.

I would like to thank my collaborators at George Washington University, including Dr. Francys Subiaul and Dr. Elizabeth Renner and at Binghamton University, including Dr. Peter Gerhardstein and Alecia Moser who have helped me learn and grow as a researcher on numerous projects over the past five years.

I would like to thank all my friends and colleagues in the Psychology Department who have provided support along the way, in particular the many graduate students who have created such an encouraging work environment to thrive in. An enormous thank you goes to my fellow 5th years, Kristin Brethel-Haurwitz, Erika Fountain, Anna Markowitz, and Nicole Senft. I cannot imagine going through this experience without your support and friendship.

Finally, and most importantly, I would like to thank my advisor Dr. Rachel Barr for her tremendous support and encouragement. You have helped cultivate my interests and I feel very fortunate to have had the opportunity to work under your mentorship. You provided me with guidance and allowed me to grow as a researcher. I truly could not have been able to undertake this dissertation without you and I feel very lucky to have worked with you on so many projects over the past five years.
I would like to dedicate this work to my family for the love I have received from them. Thank you for encouraging me to follow my passion and supporting me along the way. Last but not least, my best friend Josh – I am so lucky to have found someone as incredible as you. Thank you for your love, patience, and encouragement throughout this process.

Many thanks,
LAURA J. ZIMMERMANN
TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION ................................................................................................................. 1

CHAPTER II: EXAMINING THE IMPACT OF COGNITIVE LOAD ON WORKING MEMORY IN 18- AND 24-MONTH-OLDS: THE EFFECT OF INCREASING MEMORY DEMANDS AND REDUCING PERCEPTUAL CUES IN INFANTS AND TODELRS

Experiment 1 .............................................................................................................................................. 18
Method .................................................................................................................................................... 19
Results .................................................................................................................................................... 22
Discussion .............................................................................................................................................. 25
Experiment 2 .............................................................................................................................................. 25
Method .................................................................................................................................................... 27
Results .................................................................................................................................................... 29
Discussion .............................................................................................................................................. 31

CHAPTER III: VISUOSPATIAL WORKING MEMORY DEVELOPMENT: AN EXAMINATION OF AGE-RELATED CHANGES AND SEX DIFFERENCES IN 2- TO 4-YEAR-OLDS

Method .................................................................................................................................................... 43
Results .................................................................................................................................................... 48
Discussion .............................................................................................................................................. 57

CHAPTER IV: A NEW VISUOSPATIAL WORKING MEMORY TOUCHSCREEN TASK AND ITS’ RELATIONSHIP TO INHIBITORY CONTROL AND MOTOR-SPATIAL IMITATION IN PRESCHOOLERS

Method .................................................................................................................................................... 71
Results .................................................................................................................................................... 77
Discussion .............................................................................................................................................. 88

CHAPTER V: GENERAL DISCUSSION ................................................................................................... 93

REFERENCES .......................................................................................................................................... 101
LIST OF FIGURES

Chapter I
Figure 1. Baddeley’s Working Memory Model .......................................................... 7

Chapter II
Figure 2. Petrides’ multiple boxes task ........................................................................ 15
Figure 3. Delayed non-matching to sample ................................................................. 17
Figure 4. Warm-up Phase of Hide the Pots ................................................................. 21
Figure 5. Test Phase of Hide the Pots .......................................................................... 21
Figure 6. Performance on the HTP task (±1SE) as a function of age (18 and 24 months) delay between hiding and search (2s v. 10s) ........................................................................ 24
Figure 7. Hiding and test phase of same colored cups Hide the Pots ........................... 31

Chapter III
Figure 8. Hiding Phase and Retrieval Phase of Spin the Pots ..................................... 44
Figure 9. Success rate working memory scores (±1SE) as a function of age and sex for participants who retrieved 6 stickers ........................................................................ 52
Figure 10. Perseveration (left) and Alternate Perseveration scores (right) (±1SE) as a function of age and sex for participants who retrieved 6 stickers ........................................................................ 54

Chapter IV
Figure 11. CANTAB Spatial Working Memory Task .................................................. 66
Figure 12. Find the Stars trial configuration of varying loads and retrieval phase where stars were found ........................................................................ 73
Figure 13. Motor-Spatial Task Loads ......................................................................... 74
Figure 14. Motor-Spatial Task Test ............................................................................. 74
Figure 15. Flanker Fish sample practice trials .............................................................. 75
Figure 16. Flanker fish test trials ................................................................................. 76
Figure 17. Mean span scores by cognitive load and composite score on Find the Stars by age group ........................................................................ 80
Figure 18. Mean accuracy scores on Flanker Fish by age in years ............................. 82
Figure 19. Mean ratio scores by cognitive load and composite score on Motor-Spatial Imitation by age group ................................................................. 83
LIST OF TABLES

Chapter I
Table 1. VSWM tasks with non-human primates, adults, and children ...........................................7

Chapter II
Table 2. Correlation Table .....................................................................................................................23
Table 3. Mean working memory and perseveration scores (SDs) as a function of age and sex....24
Table 4. Correlation Table .....................................................................................................................30
Table 5. Mean working memory and perseveration scores (SDs) by cup condition .................31

Chapter III
Table 6. Mean time to complete task, success rate, adjusted working memory score, total trials (SDs)........................................................................................................................................51
Table 7. Perseveration, alternate perseveration, non-perseveration errors, and first error (SDs) as a function of age and sex of the child for those who retrieved 6 stickers.............................................53
Table 8. Correlation Table .....................................................................................................................57

Chapter IV
Table 9. Participants by Find the Stars Item Length..............................................................................79
Table 10. Partial Correlation Table for Find the Stars (FTS) composite, Motor-Spatial (MS) composite, Flanker Fish (Flanker), BRIEF-P, and demographic variables controlling for age in days ......................................................................................................................86
Table 11. Predictors of Find the Stars.....................................................................................................88
CHAPTER I: INTRODUCTION

Core executive functioning (EF) skills include working memory, inhibition, and shifting. They are inter-connected and critical skills for successful cognitive development (Diamond, Barnett, Thomas, & Munro, 2007). Working memory is defined as the manipulation, updating, and retention of information in the short term (Garon, Bryson, & Smith, 2008, Gathercole, 1999). Furthermore, Baddeley’s definition of working memory as a “temporary storage system under attentional control that underpins our capacity for complex thought” (Baddeley, 2007, p.1) highlights the crucial role of working memory for complex cognitive operations. Most consider the maintenance of information in memory different from working and is generally referred to as short-term memory. While working memory and inhibitory control are dissociable (Best & Miller 2010, Garon et al 2008), complex working memory tasks that involve maintaining, manipulating and updating a significant number of items co-activate inhibitory control (Garon et al., 2008). Inhibitory control failure is indexed via perseveration, the repeated search to a previously rewarded location (Diamond, 1990).

Working Memory Models

When Baddeley and Hitch’s (1974) working memory model was proposed four decades ago, it was described as involving three basic components termed the ‘central executive,’ ‘phonological loop,’ and ‘visuospatial sketchpad.’ The central executive is the component of the attentional-control system that ensures working memory resources are allocated adequately for some purpose or end-goal (Baddeley, 1992). The visuospatial sketchpad processes visual and spatial information and the phonological loop processes speech-based information (Baddeley, 1974). A fourth component, the episodic buffer, was proposed more recently and is thought to provide temporary storage of information from multiple modalities with the potential for
integration of perceptual cues (See Figure 1). Areas shaded in grey are systems capable of long-term storage while others represent attention and temporary storage. Before the term visuospatial working memory (VSWM) was coined it was referred to as the visuo-spatial sketchpad, which deals with all visuospatial material. More specifically, VSWM refers to the capability of working memory to store and process visual and spatial information, such as shapes. Auditory working memory, previously coined phonological loop (Baddeley, 1974) refers to the capability of working memory to store and process auditory information such as speech.

![Baddeley's Working Memory Model](image)

*Figure 1. Baddeley’s Working Memory Model. Published in Baddeley, A.D. (2000). The episodic buffer: a new component of working memory? Trends in cognitive sciences, 4:11, 417-423. Figure 1b, p. 418.*

In contrast to the Baddeley (2000) Working Memory Model (Fig. 1), Cowan proposed an embedded-process model of working memory (Cowan, 1988). This model focuses more on underlying cognitive processes when problem solving or making a decision rather than the components of working memory as Baddeley proposed. The four key elements include the *focus of attention*, *central executive* which directs attention, *activated-memory* which is analogous to Baddeley’s *episodic buffer* that stores information for a short time, and *long-term memory*. In striking contrast to Baddley’s model, Cowan’s model does not make the distinction between
processing information in different modalities. This is inconsistent with a large body of the literature demonstrating clear modality effects between auditory and visuospatial processing (See Baddeley, 2003 for review; Reuter-Lorenz, Jonides, Smith, Hartley, Miller, Marshuetz, & Koeppe, 2000).

Cowan’s (1988) focus on process has illuminated components of successful working memory performance that must be included in task design. Successful performance on a working memory task is characterized by the ability to mentally retain and manipulate information for a short time. This involves maintenance rehearsal, updating, and inhibitory control. Maintenance rehearsal refers to the ability to maintain information during a delay period when it is no longer perceptually available (D’Esposito, Postle, Ballard, & Lease., 1999; Diamond, 2013). Updating involves integrating new information into your thought process or action plan (Diamond, 2013). Inhibition involves the control of attention through cognitive engagement or disengagement about responses or rules (Harnishfeger, 1995; Nigg, 2000). In the context of VSWM tasks inhibition is conceptualized as the suppression of a prepotent response, typically a motor response such as reaching (Diamond, 1990; Espy & Bull 2005). The development of memory models in conjunction with the development of task development with non-human primates, children, and adults has informed our understanding of working memory however, gaps remain in conceptualizing these changes in early childhood.

Developmental Models of Executive Function

The relationship between the three EF components (working memory, inhibitory control, and shifting) have been explored within the context of early childhood development. Garon and colleagues proposed a hierarchical model of EF development (Garon, Bryson, & Smith, 2008). This model extends that of Miyake and colleagues (2000) “unity and diversity” theoretical
framework that suggests that EF consists of related but distinct components—working memory, inhibition, and shifting (Miyake, Friedman, Emerson, Witzki, Howarter, & Wager, 2000; Miyake & Friedman, 2012). The hierarchical model is an integrative framework that posits attention is the early foundation and that EF develops hierarchically. Based upon developmental findings, Garon suggested that early simple skills emerge next, such as delaying a response or maintaining information, in the first 3 years of life. Then simple components are integrated into complex EF processes such as working memory or inhibition. Lastly, later in development even more complex abilities emerge such as shifting, which requires multiple EF components. As Diamond summarizes, the relationship between working memory and inhibitory control is still not well understood as there is disagreement if inhibition is separate from WM, a product of exercising WM, or WM and inhibition depend on the same limited capacity system (Diamond, 2013).

This dissertation draws on the hierarchical model in its’ investigation of age-related changes in the development of VSWM across early childhood. Recent support for this model comes from findings from a novel EF battery for preschoolers highlighting age-related changes in 1.5- to 5-year-olds (Garon, Smith, & Bryson, 2014). Specifically, children performed better on simple versus complex EF measures. But prior to providing evidence in support of a theoretical framework, more basic research is needed to address gaps in our understanding of the relationship between specific factors, such as cognitive load (by varying task complexity), and working memory performance beyond infancy.

**Working memory gaps in early childhood**

The first five years are characterized by development of executive functions (EF) (Garon, Bryson, & Smith, 2008), and rapid brain growth. By age 5, children’s brains are about 90% the size of adults (Dekaban & Sadowsky, 1978). Behaviorally, working memory development has
been shown to rapidly develop during the first 5 years and to be predictive of later outcomes. In the school setting, working memory at age 5 has been suggested to better predict academic success, including reading and math, than IQ (Alloway & Alloway, 2010; Bull & Scerif, 2001) and is related to numerous positive social or contextual factors such as parent-child interaction quality (Carlson, 2009). Even later in childhood this trend persists. Better working memory between the ages of 6 and 16 is predictive of future achievements in skills such as mathematics and reading (Gathercole, Pickering, Knight, & Stegmann, 2004; Dumontheil & Klingbert, 2012).

Working memory has been extensively examined beginning during infancy using tasks such as the A-not-B paradigm (Diamond, 1990; 1997). From existing research that has been conducted to understand how working memory abilities develop in early childhood, we know young infants have difficulty inhibiting a prepotent response for a location that previously contained a hidden object, suggesting impairments in VSWM updating (Diamond, 1990). There are also age-related changes in the first year of life as older infants can succeed on A-not-B with a longer delay between a hiding and retrieval event than younger infants. (Diamond, 1990).

In toddlerhood, researchers have predominately examined EF composites in relation to other social factors, such as parent child interaction quality (Carlson, 2009) rather than purely examining changes in VSWM during the first few years of life. Additional work with preschoolers is also inconclusive as many of the tasks are confounded with verbal report measures to assess working memory (NIH Toolbox List Sorting Working Memory Test, Bauer & Zelazo, 2013) resulting in floor effects with children under 5 or suffer from technical limitations and task complexity (CANTAB Spatial Working Memory Task, Luciana & Nelson, 1998).

Table 1 highlights a review of existing VSWM tasks that have been developed for and tested with non-human primates, adults, infants, toddlers, and preschoolers. As summarized in
Table 1, while, advances in task development have been made, much of the focus and progress has been made with infants under 12 months or toddlers 3 years and up and even then our knowledge about age-related changes in working memory during the preschool years is still limited by ongoing problems with task design. As seen on page 7 less work exists with toddlers between 1- and 3-years-of-age that examines VSWM as a distinct component. The fact that tasks with toddlers and preschoolers have failed to adequately assess age-related changes in VSWM is somewhat surprising and problematic given rapid physiological and general WM developmental changes taking place leaving us with a limited understanding of how it develops.

There are several important points to note in the literature review in Table 1. While extensive research has been conducted with infants under the age of one, using variations of the A not B task, we do not have a clear understanding of how performance varies later in infancy as fewer studies have been developed and tested with 18- and 24 month-olds (see Hide the Pots-Bernier, Carlson, & Whipple; Spin the Pots- Hughes & Ensor, 2005), toddlers (see Spin the Pots-Hughes & Ensor, 2005), and preschoolers (see Corsi Block- Milner, 1971, CANTAB-Luciana & Nelson, 1998, Change Detection- Simmering, 2012). There are few age-appropriate VSWM tasks available for toddlers. Second, there are a number of tasks that have been developed to assess VSWM that have been used at discrete age-ranges and far fewer that are available for use across a wide age range that spans from infancy through the preschool years or from toddlerhood through preschool. Third, multiple factors that relate to WM performance such as directedness (child or experimenter), cognitive load, delay are not typically manipulated or compared within a task.
Table 1. VSWM tasks with non-human primates, adults, and children

<table>
<thead>
<tr>
<th>Task</th>
<th>Age</th>
<th>Directed by</th>
<th>Analogous Tasks</th>
<th>Strengths</th>
<th>Limitations/Future Directions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple Boxes (Pinto-Hamuy &amp; Linck 1965; Petrides, 1988)</td>
<td>Non-human primates with frontal lesions</td>
<td>Self</td>
<td>Self-ordered pointing, Pinto-Hamuy &amp; Linck’s task</td>
<td>Stationary version and location change version, stationary version performed by frontal lesion participants</td>
<td>N/A</td>
</tr>
<tr>
<td>Pinto-Hamuy &amp; Linck’s follow-up task, 1965</td>
<td>Non-human primates with frontal lesions</td>
<td>Self &amp; Experimenter versions</td>
<td>Self-ordered pointing, Spin the Pots, Corsi Block</td>
<td>Earliest known primate study demonstrating the frontal lobe lesions impair self-directed VSWM; internally and externally ordered versions</td>
<td>N/A</td>
</tr>
<tr>
<td>Delayed non-match to sample and Delayed match to sample (Mishkin &amp; Delacour, 1975)</td>
<td>Non-human primates</td>
<td>Experimenter</td>
<td>N/A</td>
<td>Primate study examining both match and non-match conditions</td>
<td>N/A</td>
</tr>
<tr>
<td>Multiple Wells Task (Mishkin &amp; Delacour, 1975)</td>
<td>Non-human primates</td>
<td>Experimenter</td>
<td>Multiple boxes, Spin the Pots, Hide the Pots</td>
<td>Performed with varying number of locations</td>
<td>N/A</td>
</tr>
<tr>
<td>Self-ordered pointing (Petrides &amp; Milner, 1982)</td>
<td>Adults with frontal and temporal lesions</td>
<td>Self</td>
<td>Multiple boxes, Spin the Pots, CANTAB SWM</td>
<td>One of the earliest studies to demonstrate the integral role of the frontal lobe in VSWM in humans</td>
<td>N/A</td>
</tr>
</tbody>
</table>
### Table 1. (continued)

| Task                                      | Age                  | Directed by | Analogous Tasks                           | Strengths                                                                                     | Limitations/Future Directions                                                                 |
|-------------------------------------------|----------------------|-------------|--------------------------------------------|------------------------------------------------------------------------------------------------|
| *N-back* (Kircher, 1968)                  | Adults               | Experimenter| *Change Detection*, delayed match to sample | Can be used with fMRI, multiple things can be tracked (letters, shapes, locations, etc.)         | Difficult for children to understand task, button press also has been documented to be challenging for younger participants, many adaptations possible |
| *Change detection* (Phillips, 1974)       | Adults               | Experimenter| *N-back, delayed non-match to sample*      | Also performed with children                                                                   |                                                                                               |
| *A not B* (Piaget, 1954)                  | 6-12 months          | Experimenter| Delayed Response with non-human primates   | Illustrated age-related changes, tested comparatively, extensively modified (delay, increasing number of wells, landmark or cue information), PKU, large datasets (Bell 2010; Cuevas & Bell 2011) | Possibilities exhausted                                                                       |
| *Delayed response*                        | 5 months-12 months   | Experimenter| A-not-B (identical)                        | Tested comparatively                                                                           | Possibilities exhausted                                                                       |
| *Visual Search A-not-B* (Bell & Adams, 1999) | 3- to 12-months       | Experimenter| A-not-B, Delayed response, Peek-a-boo A-not-B | Able to test younger infants as it does not require reaching                                     | Possibilities exhausted                                                                       |
| *Peek-a-boo A not B* (Reznick, Morrow, Goldman, & Snyder, 2004) | 4.5-6.5 months       | Experimenter| A not B and Delayed response but with social cues (faces) instead of objects | Social stimuli, made it possible to test younger infants with A not B paradigm due to eye gaze measurement instead of reaching | Limited age range, but could be extended through 12 months but not beyond |
Table 1. (continued)

<table>
<thead>
<tr>
<th>Task</th>
<th>Age</th>
<th>Directed by</th>
<th>Analogous Tasks</th>
<th>Strengths</th>
<th>Limitations/Future Directions</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Multiple Boxes-scrambled and stationary</em></td>
<td>6 months-7 years*</td>
<td>Experimenter</td>
<td><em>Hide the Pots, Spin the Pots, Multiple Boxes (primates), self-ordered pointing</em></td>
<td>Tested comparatively with primates, formed the basis for self-ordered pointing tasks</td>
<td>*Spans multiple age categories, vary load depending on number of boxes, delay, or spatial configuration</td>
</tr>
<tr>
<td><em>GAP-Toddlerhood</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hide the Pots (Bernier, Carlson &amp; Whipple, 2010)</td>
<td>18 months</td>
<td>Experimenter</td>
<td>Stationary boxes (cite), 3-item Multiple Boxes (Petrides, 1988), 3-item Multiple Wells, Multiple Boxes-stationary</td>
<td>3-locations, quick to administer</td>
<td>Task is too easy for 24 month-olds, include longer delay between hiding and retrieval (see Diamond, 1990)</td>
</tr>
<tr>
<td>Spin the Pots (Hughes &amp; Ensor, 2005)</td>
<td>2-year-olds</td>
<td>Self</td>
<td>7-item Multiple Wells, Multiple Boxes-scrambled</td>
<td>8-locations, children very engaged in game, takes a few minutes to administer</td>
<td>Reported as EF composite score limiting WM understanding, task is too difficult for 24-30 month-olds; modify from 8-item to 6-item (with 6 occluded objects)</td>
</tr>
<tr>
<td>Visual displacement (Morasch &amp; Bell, 2011; Cuevas &amp; Bell, 2013)</td>
<td>24- to 48-months</td>
<td>Experimenter</td>
<td>A-not-B task</td>
<td>Tested in wide age range, part of longitudinal analysis</td>
<td>N/A</td>
</tr>
<tr>
<td>Corsi Block and Reverse Corsi Block (Milner, 1971)</td>
<td>3-year-olds</td>
<td>Experimenter</td>
<td>Spatial analog to verbal WM task (digit span and reverse digit span), Multiple Boxes</td>
<td>Externally ordered sequence, wide age range tested</td>
<td>Only involves maintenance (Corsi Block), but Reverse Corsi Block involves updating in addition to maintenance (complex WM)</td>
</tr>
</tbody>
</table>
Table 1. (continued)

<table>
<thead>
<tr>
<th>Task</th>
<th>Age</th>
<th>Directed by</th>
<th>Analogous Tasks</th>
<th>Strengths</th>
<th>Limitations/Future Directions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visuospatial WM (VWM) dot matrix from Automated Working Memory Assessment (AWMA; Alloway et al., 2006)</td>
<td>4.5- to 11.5-year-olds</td>
<td>Experimenter</td>
<td>Delayed match to sample</td>
<td>Performed on computer</td>
<td>N/A</td>
</tr>
<tr>
<td>Mazes memory task (modified from WMTB-C Pickering &amp; Gathercole 2001)</td>
<td>4.5- to 11.5-year-olds</td>
<td>Experimenter</td>
<td>N/A</td>
<td>Performed on computer</td>
<td>N/A</td>
</tr>
<tr>
<td>CANTAB self-ordered searching task (Luciana &amp; Nelson, 1998; Sahakian &amp; Owen, 1992)</td>
<td>4-8 year olds</td>
<td>Self</td>
<td>Self-ordered search task (alternate name), touchscreen version of Multiple Boxes</td>
<td>Touchscreen task, valid for multiple age ranges with modifications</td>
<td>Previous findings with this age range were confounded by task complexity and technical limitations- touchscreen insensitivity; future possible modifications</td>
</tr>
<tr>
<td>Change detection task (Simmering, 2012)</td>
<td>3.5- to 4.5-year-olds</td>
<td>Experimenter</td>
<td>Change detection task with adults, delayed match to sample, dot matrix, n-back</td>
<td>Performed using NIRS</td>
<td>Requires verbal report but could be modified to make task easier for younger children</td>
</tr>
</tbody>
</table>
Based on the review of VSWM object retrieval tasks available that have been used in early childhood (see Table 1) this dissertation seeks to address this gap in the literature examining age-related changes in VSWM during early childhood. The following experiments have been conducted to examine age-related changes in VSWM from infancy through preschool and factors associated with performance.

**The Dissertation Aims**

**Aim 1: Examining the impact of cognitive load and perceptual color cues on infant working memory**

The aim of the second chapter is to examine factors associated with performance on a simple VSWM task in infancy through toddlerhood. Using an established task, *Hide the Pots* (Bernier, Carlson, & Whipple, 2010) a simple 3-location object occlusion task. Chapter 2 seeks to examine the extent to which age, delay, and perceptual color cues effect performance on a working memory task in 18- to 24 month-olds. Study 1 examines age-related changes in VSWM between 18- and 24-months of age, on *Hide the Pots* following a 2s or 10s delay. Study 2 examines the effect of varying perceptual color cues on performance in 18-month-olds.

**Aim 2: Examining age-related changes and sex differences in working memory in toddlerhood and preschool**

The aim of the third chapter is to conceptualize age-related changes in children from 2- to 4-years on an established VSWM task, *Spin the Pots* (Hughes & Ensor, 2005) a complex, 8-location object occlusion task. The study examines the extent to which age and sex are associated with performance on a working memory task in toddlers and preschoolers. Participants (*N* = 621) between the ages of 2- and 4-years were tested in the home or in a museum.
**Aim 3: Developing a new touchscreen visuospatial working memory task and examining its’ relationship to inhibitory control and spatial imitation in preschoolers**

The aim of the fourth chapter is to adapt the Cambridge Neuropsychological Test Automated Battery Spatial Working Memory (CANTAB SWM) task for children 3- to 5-years to provide a nonverbal and age appropriate measure of VSWM. We sought to examine the extent to which age and cognitive load affect performance on a new working memory task as well as it’s association to inhibitory control and spatial imitation in 3- to 5-year-olds. In a study at the National Building Museum 80 preschoolers were tested on an executive function battery. Children performed three tasks on a touchscreen: 1) *Find the Stars* (adapted from CANTAB SWM), 2) *Flanker Fish* (adapted for touchscreen from Zelazo et al., 2013), and 3) *Motor-Spatial Sequencing* (Subiaul, et al., 2012).
CHAPTER II: EXAMINING THE IMPACT OF COGNITIVE LOAD ON WORKING MEMORY IN 18- AND 24-MONTH-OLDS: THE EFFECT OF INCREASING MEMORY DEMANDS AND REDUCING PERCEPTUAL CUES

A growing body of evidence suggests there are age-related changes in VSWM within the first year of life. The examination of infant working memory capacity began six decades ago with the classic A-not-B task, a non-verbal object retrieval task (Piaget, 1954). In the A-not-B task, the experimenter hides a toy in location A and the infant retrieves it. After a predetermined number of trials at location A, the experimenter hides the toy in location B. In his seminal studies Piaget (1954) demonstrated that although 7- to 12-month-olds were consistently successful at retrieving the toy in location A, they failed to find it when the location changed to location B, the now classic A-not-B error (Piaget, 1954). Piaget (1954) attributed the A-not-B error to infants’ inability to update their object representation to account for the object’s new location. More recently however, researchers have interpreted this A-not-B error to be due to the inability of infants to inhibit a previously correct motor response for the initial location (Diamond, 1990). As we began to better understand the A-not-B error in infancy, comparative researchers were studying an identical paradigm separately with non-human primates.

Studies of non-human primates were conducted at the same time using the delayed response task. The protocol for the delayed response task was almost identical to the A-not-B task except that the alternating patterns of hiding in locations A and B were randomized. Research was conducted in parallel with humans and non-human primates using these almost identical tasks for almost 50 years until Diamond (1990) directly compared performance by 7 to 12 month-old infants on the A-not-B and delayed response tasks and found that patterns on the two tasks were the same. This allowed Diamond to compare performance by human infants to
infant rhesus monkeys on the delay response task as well. Across species delay has been shown to negatively impact performance on the A-not-B task.

**Age-related changes**

Human infants between 7.5 and 12 months showed a linear increase in performance on the A-not-B and delayed response tasks. In addition, Harris (1973) reported that 9- to 10-month-olds performed significantly better with no delay compared to a 5 second delay. In a more extensive series of studies, Diamond found that although 9-month-olds could perform the task after a 4 second delay, older 12-month-olds infants and rhesus monkeys could perform the task after a much longer 10 second delay (see Diamond, 1990; Diamond, Prevor, Callender, & Druin, 1997). Together, these results highlight age-related performance improvements in VSWM with infants up to 12 months-of-age using reaching behavior measures as well as manipulating delay.

These age-related changes may be due to the rapid development of the prefrontal cortex during infancy. Studies with both human and non-human primates have indicated a role of prefrontal cortex during VSWM tasks. In human infants, later studies using the A-not-B task while measuring neural activity demonstrated age-related increases in EEG power and coherence between 5 and 10 months (Cuevas & Bell, 2011). The authors attributed this finding to increased localized activation and connectivity in the frontal lobes. Success on the delayed response task in rhesus monkeys has also been linked to regions involved in the memory network including the dorsolateral prefrontal cortex (DLPFC) (Diamond & Goldman-Rakic, 1989) which is rapidly developing during early childhood.

Although we have a comprehensive understanding of VSWM indexed via the A-not-B and delayed response tasks during the first 18 months of life, much less is known about the development of VSWM from late infancy to early childhood. Researchers turned to work
conducted with non-human primates in order to develop more complex VSTM tasks to assess older infants and toddlers.

**Non-human primate studies**

The earliest VSTM tasks with non-human primates laid the groundwork for task development and conceptualizing VSTM development in early childhood. Pinto-Hamuy and Linck (1965) first developed the *multiple boxes task*. The task is to retrieve a reward from under three different shaped and colored boxes while not repeating a search to a previously rewarded location. During the *multiple boxes task* objects were hidden under 3 colorful pots (Figure 2). On each trial monkeys pointed to one pot to retrieve the object. After each trial an occluder was raised and the experimenter scrambled the boxes. On the next trial the monkey then attempted to locate another pot that contained food. On this task monkeys must encode color as a primary cue, rather than spatial location. Petrides (1988) selectively lesioned either the hippocampus or the DLPFC in rhesus monkeys. This self-directed task recruited the DLPFC because performance on the multiple boxes task was impaired in monkeys with DLPFC lesions but intact in monkeys with hippocampal lesions.

*Figure 2. Petrides’ multiple boxes task. (Image Published in Gluck, M.A., Mercado, E. & Myers, C.E. (2008). Learning and Memory: From Brain to Behavior. USA: Worth Publisher, p.179).*
In a follow-up study, Pinto-Hamuy and Linck (1965) investigated the impact of frontal lesions on internally ordered and externally ordered WM tasks in rhesus monkeys. The task consisted of square-shaped cued panels presented on a computer screen using an automated discrimination apparatus ranging from 1-12 colored stimuli on 1.5x2.5 inch Lucite panels. The pattern was preprogrammed. Sample sequences included red-red-red or green-orange. In the internally ordered task monkeys had to push on the series of panels without repetition as they had done in the multiple boxes described above. This is in contrast to the externally ordered task in which monkeys had to push on a series of panels determined by external cues. Researchers demonstrated that monkeys with frontal lesions were impaired on the internally ordered task, but they performed no differently than monkeys without lesions when external cues were provided (Pinto-Hamuy & Linck, 1965).

Mishkin & Delacour (1975) developed the delayed non-match to sample and delayed match to sample task (Figure 3). In both cases monkeys were shown a novel object under which was hidden a piece of food. Following a delay period on the scale or seconds to minutes, monkeys searched for the match (location that contains same novel object) or non-match (location that contains different novel object) to retrieve the reward. To succeed on the task, monkeys must hold the initial configuration in mind during the delay until presented with the new configuration and then update their memory for the correct object.
Stationary boxes task. Taking advantage of work conducted with non-human primates, more complex tasks were developed and tested with older infants and toddlers. For example, borrowing components from the multiple boxes task in the stationary boxes task, children view a stationary array of boxes and have to search for a reward within each location per trial. The experimenter first hides a toy in all locations. A delay is imposed between trials and participants must try to search in a new box each time. That is, they must update their memory after each trial to avoid searching in an empty box. Diamond, Prevor, Callender, and Druin (1997) observed age-related changes on this task. In the 15- to 30-month group (3 boxes): toddlers needed fewer reaches to open all boxes and perseverated less with age. Similar results have also been demonstrated with a 10- to 30-month cohort on the 3 boxes task (Ewing-Cobbs, Prasad, Landry, Kramer, & DeLeon, 2004). The stationary box task involves self-directed search that may be more cognitively challenging for children than finding a hidden objects after watching an experimenter hide them (see Chapter 3 of this dissertation).

Together these research findings suggest that a short delay can negatively impair performance for infants under one year of age. While VSWM research in toddlerhood is limited,
perhaps in part due to a task gap (see Chapter 1, Table 1), some research suggests there are age-related changes in early childhood on object retrieval in the first few years of life. The present study seeks to better understand these changes in early childhood using a simple VSTM task.

The Present Study

To better understand factors associated with the ability to update and manipulate visuospatial information in early childhood we used the *Hide the Pots* (HTP) VSTM task. The HTP task was developed by Bernier, Carlson, & Whipple, 2010. It is an experimenter-directed ordered stationary multiple boxes task. The experimenter hides an object under 1 of 3 cups and closes the lid of a box for 1 to 2 seconds at which time the child is allowed to retrieve the hidden object. On trials 2 and 3, the experimenter hides the object in a new location. In HTP it is possible to use either color or location as a cue to the hidden object but participants must encode and rehearse the experimenter’s hiding destination, which changes location arbitrarily three times. Researchers using HTP have not examined age-related changes in VSTM performance or manipulated task parameters such as delay or color.

Experiment 1. Age-related changes and the impact of delay in working memory

The aim of Experiment 1 was to examine whether increasing cognitive load factors via imposing a longer delay between hiding and search, would affect VSTM from 18- to 24-months using HTP. In Experiment 1, independent groups were tested after a 2s or 10s delay between hiding and search. Groups of 18- and 24-month-olds were compared to examine age-related changes. It was hypothesized that 24-month-olds would be more efficient at solving the task (better working memory score and fewer perseverations) than 18-month-olds. In addition, it was
further hypothesized that increasing the delay between hiding and retrieval would impair working memory performance.

**Method**

**Participants**

Data were collected from 231 typically developing children (108 boys) at 18 and 24 months in their homes in the greater Washington, DC area. Portions of data collected with these participants are published or under review elsewhere (Barr, Brito, Nugent, & Zimmermann, under review; Barr, Moser, Rusnak, Zimmermann, Dickerson, Lee, & Gerhardstein, under review; Zimmermann, Moser, Grenell, Dickerson, Yao, Gerhardstein, & Barr, 2015). Independent groups of children were tested at 18 months ($N = 62, M$ age = 18 months 18.24 days, $SD = 48.92$ days) and 24 months ($N=169, M$ age = 24 months 18.72 days, $SD = 42.18$ days). Participants were primarily Caucasian (77.1%) and from graduate school-educated families ($M$ years of education = 17.55, $SD = 1.14$). The remaining sample included the following races: mixed (15.2%), African-American (3%), Asian (2.2%), other (1.3%), and not reported (1.3%). With regard to ethnicity, 2.6% of the sample was Latino. Less than half of children lived in homes where more than one language was routinely spoken (38.5%). The mean rank of socioeconomic index (SEI) (Nakao & Treas, 1992) was 78.66 ($SD = 11.80$) based on reports by 185 families (80% reporting). Productive vocabulary measured by the MacArthur Communicative Developmental Inventory (MCDI) at 18 months (raw score $M= 25.94, SD = 19.08$) and 24 months (raw score $M= 55.44, SD = 23.44$). Of the 231 children, 18 children were excluded from the analysis: 9 due to failure to complete the task, 4 for technical error with video, 3 due to parental or sibling interference, 1 due to experimenter error, and 1 for multiple cup lifts on test trials.
Apparatus and Stimuli

The *Hide the Pots* (Bernier, Whipple, & Carlson, 2010) apparatus was used. Three distinctly colored opaque cups (red, blue, and green), a small black and white ball, and a box were used for this task. All three cups fit inside the box in a straight line with equal spacing between them and a hinge attached a lid to the box in order to easily open and close the box (see Figure 4). Three distinctly colored cups (red, green, and blue) were fit under a box, which was closed at one point during each of the three trials. The box was made of a black cardboard material measuring 12 cm tall, 25 cm long, and 17.5 cm wide.

Self-report Measure

The caregiver was asked to complete a general information questionnaire (assessing rank Socioeconomic Status (SES), parental education, and language exposure) as well as the MacArthur Communicative Development Inventory: Words and Sentences Short Form (MCDI) to measure children’s productive vocabulary (Fenson et al., 2000).

Design and Procedure

Testing occurred in the home. After obtaining informed consent, primary caregivers were asked to complete the general information questionnaire.

All participants had a warm-up and a test phase with *Hide the Pots* (HTP). The warm-up consisted of three hiding events without occlusion. The warm-up phase familiarized the child with the apparatus and expectations for subsequent trials.
During the practice trials, the infant watched as the experimenter placed a small ball under one of the three cups. The experimenter then encouraged the infant to retrieve the ball by saying, “Can you get the ball?” Once the infant retrieved the ball, the experimenter praised the infant then placed the ball under a different cup. There were a total of three practice trials so that the infant understood the rules of the task.
Children were assigned to either a 2s or 10s delay condition for the test phase. The test trials were identical to the practice trials, except that after the experimenter placed the ball under one of the cups, the box was closed for 2s or 10s. After the delay, the experimenter opened the box and once again encouraged the infant to retrieve the ball with the same verbal cue. Each trial required the infant to hold the location of the ball in memory and each subsequent trial required the infant to update his/her memory of the new location. Like the practice trials, there were a total of three test trials.

Coding

Task performance was videotaped for subsequent coding. For the HTP task, each child was given a working memory score and a perseveration score.

**Working memory.** Each infant was given a score between 0-3 based on the number of trials in which the infant selected the correct cup on the first search attempt.

**Perseveration.** Each infant was given a score between 0-2 based on the number of trials in which the infant chose the cup that was selected on the previous trial.

**Reliability.** A second independent coder scored 33.8% of the videos to determine reliability of the ratings; there was an inter-rater reliability kappa of 0.94.

Results

Preliminary analyses including bilingual status, sex of the child, SES, and parental education were conducted for each of the dependent measures (working memory and perseveration). No significant (p<.05) effects or interactions involving bilingual status, or parental education emerged in the preliminary analyses and data were collapsed across these variables in further analyses. The correlation table below shows the relationships between the
variables. Given that MCDI raw scores were correlated with the working memory score, MCDI was added as a covariate into subsequent analyses.

Table 2. Correlation Table

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Working Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Perseveration</td>
<td>-.58**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Age in days</td>
<td>0.09</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Sex</td>
<td>0.08</td>
<td>-0.06</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. SES</td>
<td>0.07</td>
<td>-0.04</td>
<td>0.08</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Parental education</td>
<td>0.08</td>
<td>-0.06</td>
<td>0.05</td>
<td>-0.11</td>
<td>.52**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. MCDI</td>
<td>.24**</td>
<td>-0.03</td>
<td>.44**</td>
<td>.16*</td>
<td>0.06</td>
<td>.21**</td>
<td></td>
</tr>
<tr>
<td>8. Bilingual</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.001</td>
<td>0.06</td>
<td>0.08</td>
<td>-0.05</td>
<td>-0.09</td>
</tr>
</tbody>
</table>

*p < .10, *p < .05, **p < .01

**Working Memory**

Because working memory also was significantly correlated with MCDI (see Table 2) it was added as a covariate to the ANOVA. To examine the effect of age and delay on working memory a 2(age, 18, 24) x 2(delay, 2s, 10s) ANOVA, with MCDI as a covariate on HTP scores, yielded a significant main effect of MCDI $F(1, 227) = 6.75, p < .01, \eta^2 = .41$, no main effect of age $F(1, 233) < 1, ns$, a significant main effect of delay, $F(1, 227) = 43.03, p < .001, \eta^2 = .16$ but no significant interaction between age and delay $F(1, 227) < 1, ns$.

**Perseveration**

A 2(age, 18, 24) x 2(delay, 2s, 10s) ANOVA on HTP scores with MCDI as a covariate on HTP scores did not yield a significant main effect of MCDI $F(1, 223) < 1, ns$, no main effect of
age $F(1, 223) < 1, ns$, a significant main effect of delay, $F(1, 223) = 12.54, p < .001, \eta^2 = .05$ but no significant interaction between age and delay $F(1, 223) < 1, ns$.

**Table 3. Mean working memory and perseveration score (SDs) as a function of age and sex**

<table>
<thead>
<tr>
<th>Age (mo)</th>
<th>Sex</th>
<th>N</th>
<th>Working Memory</th>
<th>Perseveration</th>
<th>N</th>
<th>Working Memory</th>
<th>Perseveration</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>Male</td>
<td>25</td>
<td>2.04 (.89)</td>
<td>.40 (.65)</td>
<td>3</td>
<td>1.67 (.58)</td>
<td>.67 (1.16)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>25</td>
<td>2.08 (.76)</td>
<td>.32 (.63)</td>
<td>9</td>
<td>1.0 (.70)</td>
<td>.67 (.65)</td>
</tr>
<tr>
<td>Total</td>
<td>-</td>
<td>50</td>
<td>2.06 (.82)</td>
<td>.36 (.63)</td>
<td>12</td>
<td>1.17 (.72)</td>
<td>.67 (.65)</td>
</tr>
<tr>
<td>24</td>
<td>Male</td>
<td>55</td>
<td>2.29 (.88)</td>
<td>.29 (.60)</td>
<td>25</td>
<td>1.32 (.95)</td>
<td>.72 (.84)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>64</td>
<td>2.59 (.61)</td>
<td>.20 (.41)</td>
<td>25</td>
<td>1.44 (1.08)</td>
<td>.64 (.70)</td>
</tr>
<tr>
<td>Total</td>
<td>-</td>
<td>119</td>
<td>2.45 (.75)</td>
<td>.24 (.50)</td>
<td>50</td>
<td>1.38 (1.00)</td>
<td>.68 (.77)</td>
</tr>
</tbody>
</table>

*Figure 6. Performance on the HTP task (±1SE) as a function of age (18 and 24 months) delay between hiding and search (2s v. 10s).*
Discussion

Experiment 1 findings indicate important limits on VSWM in infancy. Results did not support age-related changes on the HTP task at both 2 and 10 second delays as was predicted. Rather than age, MCDI was a better predictor of HTP suggesting productive vocabulary may be a better predictor of individual differences. Therefore, it is important to consider vocabulary, even with non-verbal VSWM tasks. This extends prior longitudinal work that has shown a link between early language skills at 2 years and later cognitive outcomes at age 8 (e.g. IQ and sequential working memory task) (Marchman & Fernald, 2008). Further work has demonstrated a link between vocabulary development and phonological working memory (Gathercole & Baddeley, 1993). This study suggests vocabulary may be important to examine in the context of non-verbal working memory tasks as well, especially in the early years. Consistent with prior research examining the impact of delay in early infancy, performance on the 2 second delay significantly exceeded that of the 10 second delay. The overall pattern of results shows changes in VSWM as a function of delay and MCDI.

Experiment 2. How do perceptual color cues affect working memory at 18 months?

Perceptual cues

Researchers have manipulated spatial and perceptual cues on object occlusion tasks with infants. For example, 1-year-old infants can perform successfully when a landmark cue is in the same location as a toy and perform less successfully when the landmark cue is adjacent to the target location (Bushnell, Lawrence, & Connell, 1995). It is not clear from the current research on VSWM development whether children are in fact using spatial cues or additional perceptual color cues to guide their search behavior. Some studies have examined whether infants can use
color cues to guide behavior. By age 2, children can sort and match objects based on color regardless of whether they comprehend color words suggesting they may have an early conceptual understanding (Soja, 1994) and later by around 3 years children prefer color to shape when examining geometric images (Mekman, Koriat, & Pardo, 1976).

As children age it has been demonstrated that there is a developmental shift from relying on location to other visual cues such as color and size (Daehler, Bukatko, Benson, & Myers, 1976). Daehler and colleagues presented 18-, 24-, 30-, and 36-month-olds with a four-choice delayed reaction task to measure memory for a hidden object. Cues varied across conditions to include size (4 different sized white boxes), color (4 same sized boxes of varying colors), and a control (4 identical sized white boxes). Following the four test trials, each of which had the three different set of boxes, conflict trials were presented. In this situation, after the hiding event, the experimenter switched the location of the two boxes so color and location were dissociable. They found children across ages performed best in the size condition, followed by the color cue condition, and lastly the control condition. On the conflict trials, when the perceptual color or size cue was redundant with location (because the boxes were switched), children across ages performed significantly better. It is possible that the conflict switch event performed by the experimenter confused younger infants or that 4-items may be a high load for the 18- and 24-month-olds to keep track of. The present study, therefore considerably simplifies the design by reducing the number of locations, keeping location constant, and directly comparing groups with and without color cues. Given that there were no age-related changes in performance in Experiment 1, Experiment 2 included only 18-month-olds. In Experiment 2, the question of whether 18-month-old infants could solve the HTP task in the absence of salient color cues was examined. It was hypothesized that removal of these cues would disrupt performance.
**Method**

**Participants**

Data were collected from 46 typically developing children (24 boys) 18 month-old children in their homes in the greater Washington, DC area and in a laboratory in Barcelona, Spain. Portions of data collected with these participants are published elsewhere (Brito, Sebastian-Galles, & Barr 2014) for 23 infants. Children were tested at 18 months ($N = 46$, $M$ age = 18 months, $SD = 1.85$ days). Participants were primarily Caucasian (89.1%) and from college-educated families ($M$ years of education = 16.63, $SD = 1.51$ years). The remaining 14% of the sample included the following races: mixed (6.5%), Asian (2.2%), other (2.2%), 10.9% of the sample was Latino. Almost half of children lived in homes where more than one language was routinely spoken (45.7%). The mean rank of socioeconomic index (SEI) (Nakao & Treas, 1992) was 73.37 ($SD = 16.29$) based on 35 families (76%). Productive vocabulary measured by the MacArthur Communicative Developmental Inventory (MCDI) at 18 months ($M = 18.75$, $SD = 15.63$). 6 children were excluded from the analysis: 3 due to multiple cup lifts on test trials, 2 due to failure to complete the task, and 1 due to experimenter error.

**Apparatus and Stimuli**

The same apparatus from Experiment 1 was used. In addition, three of the same colored cups were used for a modification to the HTP.

**Design and Procedure**

Testing occurred primarily in the home ($n = 23$ were tested in a laboratory in Spain). After obtaining informed consent, primary caregivers were asked to complete the same general information questionnaire as in Experiment 1.
This experiment utilized a within-subjects design and participants were tested on both sets of stimuli. There were two training phases and two test phases, one with the same colored cups (ex. red, red, red) and the other with the original set of different colored cups (ex. red, green, and blue). The order of the conditions was counterbalanced across participants. All participants had a warm-up and a test phase with HTP. The warm-up consisted of three hiding events without occlusion or delay. The warm-up phase familiarized the child with the apparatus and expectations for subsequent trials.

*Same-colored Hide the Pots Task*

This task is identical to the original task except that three identical colored cups (ex. red, red, red) were used.

![Same colored cups hiding phase](image1)
![Same colored cups retrieval phase](image2)

*Figure 7. Hiding and test phase of same colored cups Hide the Pots.*

**Coding**

The same measures were used for Experiment 2, however scores were calculated for each of the two test phases (same and different color cups).
Reliability. A second independent coder scored 15.2% of the videos to determine reliability of the ratings; coders recorded the color of the cup selected per trial and whether or not a ball was retrieved. Inter-rater reliability on HTP score was calculated (kappa = 0.99).

Results
Preliminary analyses including sex of the child, bilingual status, SES, and parental education were conducted for each of the dependent measures (working memory and perseveration). Parental education was associated with HTP performance on the same task. On closer examination of the data it was found that only one parent reported high school level education and when the correlations were rerun without this one participant the associations were no longer significant. We therefore conclude that these correlations were spurious and data were collapsed across parental education. No significant (p<.05) effects or interactions involving bilingual status emerged in the preliminary analyses and data were collapsed across these variables in further analyses.
Table 4. Correlation Table

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. HTP diff</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2. HTP same</td>
<td>.31*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3. Perseveration diff</td>
<td>-.55*</td>
<td>-.09</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4. Perseveration same</td>
<td>-.20</td>
<td>-.46**</td>
<td>.09</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5. Sex</td>
<td>.15</td>
<td>.19</td>
<td>.16</td>
<td>.02</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6. SES</td>
<td>.16</td>
<td>-.32</td>
<td>.02</td>
<td>.15</td>
<td>.19</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7. Parental Ed</td>
<td>.11</td>
<td>-.31*</td>
<td>.08</td>
<td>.02</td>
<td>.08</td>
<td>.52**</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8. MCDI</td>
<td>.18</td>
<td>.14</td>
<td>-.05</td>
<td>-.09</td>
<td>-.01</td>
<td>.06</td>
<td>-.22</td>
<td>-</td>
</tr>
<tr>
<td>9. Bilingual</td>
<td>.09</td>
<td>-.06</td>
<td>-.001</td>
<td>.02</td>
<td>-.05</td>
<td>.009</td>
<td>.02</td>
<td>.14</td>
</tr>
</tbody>
</table>

†p < .10, *p < .05, **p < .01

Working Memory

A repeated measures ANOVA with a Greenhouse-Geisser correction determined that there was no main effect of color cue on HTP performance, F(1,44) < 1, ns, no main effect of order of task presentation F(1, 44) = 1.50, p = .22, and no interaction between color cue and test order F(1, 44) < 1, p = ns. Overall there was no effect of color cue on performance by 18-month-olds in the HTP task.

Perseveration

A repeated measures ANOVA with a Greenhouse-Geisser correction determined that there was no main effect of color cue on perseverations, F(1,44) < 1, p = ns, no main effect of order of task presentation F(1, 44) < 1, p = ns, and no interaction between color cue and test order F(1, 44) < 1, p = ns. Overall there was no effect of color cue on performance by 18-month-
olds in the HTP task. Given MCDI was associated with performance in Experiment 1, the repeated measures design was subsequently calculated with MCDI as a covariate but was not significant and not included into subsequent interactions for either working memory or perseveration scores.

### Table 5. Mean working memory and perseveration scores (SDs) by cup condition

<table>
<thead>
<tr>
<th></th>
<th>Different</th>
<th>Same</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working Memory</td>
<td>1.65 (.77)</td>
<td>1.57 (.83)</td>
</tr>
<tr>
<td>Perseveration</td>
<td>.59 (.69)</td>
<td>.52 (.69)</td>
</tr>
</tbody>
</table>

**Discussion**

The present studies addressed a gap in our understanding of factors associated with performance on an established working memory task (*Hide the Pots*) in an under-investigated age range. By manipulating both memory and perceptual cues we were able to explore the role that these cues have on the developing VSWM memory system from infancy to toddlerhood. Contrary to our prediction, there were no significant age-related changes 18 and 24 months. Rather, MCDI was significantly associated with better performance on this task. This suggests individual differences on a non-verbal VSWM may be linked to early productive vocabulary skills. The low performance at 18 and 24 months suggests that even small delays can have significant impacts on performance in infancy through toddlerhood. Lastly, removing perceptual color cues did not disrupt HTP performance at 18 months. This suggests they are able to encode the location information in the absence of color cues and succeed on the task.
The delay findings that performance at 10s was significantly worse than 2s is consistent with previous studies in infancy (Diamond, 1990) but extends our understanding of VSWM beyond the first year of life. Importantly, even at 24 months a short delay on the scale of seconds can disrupt performance. While testing, this disruption was also observed behaviorally. Many children were eager to retrieve the ball during the longer delay; trying to open the box while it was held closed by the experimenter.

While the lack of age-related changes was unexpected, the present study illustrates there is a benefit to collecting other measures, such as productive vocabulary, that may be more sensitive than age category alone. Although not predicted, the association between VSWM and vocabulary is not surprising. This is primarily because as an infant or toddler is acquiring language their ability to learn and produce new words is linked to one’s ability to maintain and manipulate information in the short term. This finding fits with Baddeley and Hitch’s working memory model (Baddeley & Hitch 1974) as Baddeley states “working memory is a temporary storage system that underpins our capacity for thinking, it is clearly the case that it should have implications for language processing” (Baddeley, 2003 p. 203) and suggests the visuo-spatial sketchpad may be involved in maintaining representations.

Considering the above there is overwhelming evidence that early vocabulary is predictive of better cognitive development. This association between vocabulary may be important to consider in the context of future studies or early interventions given the word-gap findings particularly that SES differences in vocabulary are evident at 18-months (Fernald, Marchman, & Weisleder, 2013).

While there was a relatively large sample with bilingual groups in both studies there was a lack of bilingual differences. The limited research examining bilingual differences in non-
verbal visuospatial working memory with children under three makes it difficult to contextualize this finding and suggests this area provides a rich area future research. There was also a lack of sex differences in this age range, which is consistent with the initial HTP study that did not find a significant gender difference at 18 months (Bernier, Carlson, & Whipple, 2010).

The color cue finding was counter to the hypothesis because color is a salient cue in early childhood. During HTP, the spatial location remains constant and it is not necessary for children to encode the color cue in order to retrieve the reward. Perhaps differences in older kids may be expected, especially in the case of a more complex working memory task when spatial location cues are not held constant. On a more complex four-choice delayed reaction test, 18-month-olds performed better on retrieving the object when visual perceptual cues (e.g. color or size) were redundant with location cues were present (Daehler et al., 1976). However, results from experiment 2 suggest that performance did not vary as a function of perceptual color cues at 18-months on a the HTP task where the location of the cups remained constant.

There are two main limitations of the present studies. First, the HTP task score ranges from 0-3. The limited range makes it difficult to identify variations among children. Ideally, a future task would include additional test trials or responses to better capture VSWM. Second, the sample lacks diversity parental education is high. This limits the generalizability of the results to the broader population. Additional work should also be done to develop approaches to recruit and test participants from a more diverse population.

There is a need to develop different methodologies for studying infants. One potential way to bridge the toddler gap is incorporating the use of eye-tracking methodologies across early childhood. Recently, novel paradigms have been developed to test visual working memory (VWM) of infants. Specifically, anticipatory gaze responses using eye-tracking allow researchers
to understand predictions of object locations based on memory. Kaldy and colleagues (2015) developed a task that inverts the traditional delayed match retrieval paradigm by showing infants two different stimuli (e.g. yellow star, and blue moon), occluding them, presenting a delay, showing a third stimuli (e.g. blue moon), and then providing infants a chance to search for the hidden image that matches much like the card game Memory (O’Grady, Guillory, Blaser, & Kaldy, 2015). A reward structure of an animation is presented to infants based on a correct look to the matching hidden stimuli. Results indicated 8-month-olds performed at chance while 10-month-olds were able to hold the object-location information in working memory and thus performed significantly above chance (Kaldy, Guillory, & Blaser, 2015). This could be adapted further to look at VSWM as well by incorporating a spatial location search element to the game.

Prior research findings with HTP suggests performance is predictive of other EF dimensions later in childhood, such as inhibitory control and other social factors (Bernier, Carlson, & Whipple, 2010). Thus it may be useful to separately consider assessing inhibitory control or integrating social information within the context of a simple working memory task. Some research integrating social cues into a simple working memory task has indicated that it can facilitate performance in younger infants. Specifically, the modified peek-a-boo A-not-B task has revealed that under some conditions infants may be able to pass the A-not-B task even earlier, between 4.5 and 6.5 months (Reznick, Morrow, Goldman, & Snyder, 2004). Reznick and colleagues (2004) demonstrated a significant correlation between age and infant gaze accuracy (to the location the experimenter was hiding in) following a 1- to 2-second delay. Furthermore, gaze accuracy improved following three trials that consisted of the same hiding location. Starting as early as 4.5 months infants’ eye-gaze patterns indicate they are able to predict the experimenter’s correct location in the A-not-B peek-a-boo looking measure (Reznick et al.,
2004). Perhaps incorporating social cues may improve performance of infants and toddlers during search tasks despite high cognitive load situations such as delay as it may be more salient or rewarding.

This study helps fill a gap in the literature in examining factors associated with VSWM in an understudied age range. These findings make an important contribution to understanding the relationship between delay, perceptual color cues, productive vocabulary, and VSWM. Here a short 10 second delay significantly disrupts HTP performance at 24 months highlighting the still rapidly developing working memory system. Furthermore, productive vocabulary predicts successful performance on this task suggesting the importance of understanding verbal and non-verbal cognitive development concurrently.
CHAPTER III: VISUOSPATIAL WORKING MEMORY DEVELOPMENT: AN EXAMINATION OF AGE-RELATED CHANGES AND SEX DIFFERENCES IN 2- TO 4-YEAR-OLDS

This chapter is currently in revision: Zimmermann, L., Frank, H., Subiaul, F., & Barr, R. (in revision). Visuospatial Working Memory Development: An Examination of Age-related Changes and Sex Differences in Preschoolers, Developmental Psychobiology

Working memory is a key component of a larger executive functioning (EF) system that also includes inhibitory control and cognitive flexibility (Miyake, Friedman, Witzki, Howarter, & Wager, 2000). Impairments in this system result in difficulties selecting, maintaining, updating, and rerouting information processing (Shimamura, 2000). Working memory is defined as the ability mentally retain, update, and/or manipulate information for a short time, on the scale of seconds (Atkinson & Shiffrin, 1971). Most consider the maintenance of information in memory different from the manipulation and updating of information in memory. The former is generally referred to as short-term memory. The latter is generally considered to be working memory (Garon et al., 2008, Gathercole, 1999). While working memory and inhibitory control are dissociable (Best & Miller, 2010, Garon et al, 2008), complex working memory tasks that involve maintaining, manipulating, and updating multiple items co-activate working memory and inhibitory control (Garon, Bryson, & Smith, 2008). Inhibitory control failure is indexed via perseveration, the repeated search to a previously rewarded location (Diamond, 1990).

Baddeley’s (Baddeley & Hitch, 1974, Baddeley, 1986, Baddeley, 2007, Baddeley, 2012) highly influential model posits that working memory or WM includes separate maintenance systems that are domain specific and relatively independent. These include the “phonological loop” (auditory storage) and the “visuo-spatial sketchpad” (visuo-spatial storage). Consequently, researchers have generally measured working memory using both verbal and non-verbal tasks.
Performance on verbal and non-verbal working memory tasks, improves dramatically during the preschool years (Garon et al., 2008). However, changes in working memory continue well into adolescence and are especially evident on verbal working memory tasks (Best & Miller, 2010).

WM has been extensively examined across development. Researchers have documented that item recall increases between 3- to 5-years on verbal measures of WM using digit and word span tasks (Espy & Bull, 2005). Verbal WM measures at age 5 have been shown to be a better predictor of academic success in reading and math than IQ (Alloway & Alloway, 2010; Bull & Scerif, 2001). Between the ages of 6 and 16, better WM continues to predict future achievements in mathematics and reading (Gathercole, Pickering, Knight, & Stegmann, 2004; Dumontheil & Klingbert, 2012). Beyond academic success, WM and inhibition have also been associated with social and emotional development (Bernier, Carlson, & Whipple, 2010). Finally, WM has been shown to predict health, wealth and public safety in adults (Biederman, Monuteaux, Mick, Wilens, Fontanella, Poetzl, et al., 2006; Moffit, Arseneault, Belsky, Dickson, Hancox, Harrington, et al., 2011; Wilens, Martelon, Fried, Petty, Bateman, & Biederman, 2011).

WM tasks used with toddlers and preschoolers have a critical limitation: an overreliance on verbal responses. These verbal WM tasks are confounded by the complexity of verbal report required to complete the tasks (e.g., NIH Toolbox List Sorting Working Memory Test (LSWM), Bauer & Zelazo, 2013) often resulting in floor effects with children under 5. Specifically, the LSWM test of the NIH Toolbox requires complex verbal recall and performance is notably poor in 3- to 5-year-olds (Riggins, 2013). Other tasks suffer from technical limitations and task complexity (e.g., Cambridge Neuropsychological Test Automated Battery (CANTAB) Spatial Working Memory Task, Luciana & Nelson, 1998). Although performance improved on items of the CANTAB spatial WM task in children between 4- and 8-years, 4-years-olds performed poorly.
Given these limitations, we currently lack a cohesive and coherent understanding of the development of WM in young children.

One way to reduce verbal demands is to examine non-verbal visuospatial WM (VSWM). VSWM is comprised of visual memory, our capacity to remember what we see: shapes, colors, or features of stimuli and spatial memory, our capacity to remember information about locations and movement. Thus, visuospatial working memory (VSWM) refers to contexts in which both visual and spatial information are integrated to facilitate recall and memory in the short term. The non-verbal nature of visuospatial WM (VSWM) tasks makes them more developmentally appropriate for young children who are not yet proficient in verbal language. Typically, the task involves finding a rewarding occluded object, measured via reaching. A growing body of evidence suggests that there are age-related changes in VSWM (Diamond, 1990, Diamond, Prevor, Callender, & Druin, 1997, Ewing-Cobbs, Prasad, Landry, Kramer, & DeLeon, 2004, Johansson, Marciszko, Brocki, & Bohlin, 2015, Overman, Pate, Moore, & Peuster, 1996). In order to successfully complete a VSWM task, participants must attend to, manipulate, and update both perceptual and spatial cues. Although perceptual and spatial cues have been manipulated in prior research, their role in the development of VSWM during early childhood remains poorly understood.

**Age Differences in VSWM Performance**

Some tasks require greater attention to spatial cues than to individual items. In the *stationary boxes* task, children view a stationary array of identical boxes and have to search for a reward hidden under each of the locations, with one search attempt of their choice per trial. A delay is imposed between trials and participants must try to search in a new box each time. Successful completion of the task involves storing location-specific information and inhibiting
responding to previously rewarded locations while maintaining and updating spatial location. For example, in one study, Ewing and colleagues used the stationary boxes task (Ewing-Cobbs, et al., 2004) to assess age-related changes in VSWM in three groups of children from 10- to 85-months-of-age. The youngest children received a three-box array (10-30 months), preschoolers received six-box array (31-59 months), and elementary school age children received a nine-box array (> 60 months). There were age-related changes in the typically developing group demonstrated by increases in efficiency of search (indexed by the number of correct boxes/number of reaches) and a decrease in perseveration errors (indexed by the number of reaches to a previously correct location). The decrease in perseveration errors is thought to index an increase in inhibitory control.

Other VSWM tasks require greater attention to perceptual cues. In contrast to the stationary boxes task described earlier, in the scrambled boxes task, children view an array of different boxes and have to search for a reward within each location per trial; however, the location of the boxes changes after each trial. In the scrambled boxes task the child must attend to specific perceptual features of the boxes, such as the color, rather than the spatial location. The scrambled boxes task also involves inhibitory control but in this case, to a previously rewarded item rather than a previously rewarded spatial location. Diamond, Prevor, Callender, and Druin (1997) directly compared performance on the scrambled and stationary boxes tasks in a group of typically developing children and those who had PKU, a metabolic disorder that results in frontal lobe impairments. Given the hypothesized centrality of the frontal lobes in WM (Diamond, & Goldman-Rakic, 1989), this latter group was predicted to evidence delayed or atypical WM development in comparison to normal controls. Diamond and colleagues (1997) reported age-related changes on both tasks. In the 15- to 30-month group (3 boxes task, both on the scrambled
and stationary tasks), older toddlers needed fewer reaches to open all boxes and fewer reaches to the same location than younger toddlers. In the 3.5- to 7-year-old group (6 boxes), performance in the stationary boxes was better than scrambled boxes and improved with age in both stationary and scrambled conditions (Diamond et al., 1997).

In contrast to the stationary and scrambled box tasks that isolate spatial and visual cues, respectively, numerous tasks require attention (and integration) of both spatial and visual cues. The frequently used radial arm maze task that was developed for testing VSWM in rodents has been adapted for use with children (Mandolesi, Petrosini, Menghini, Addona, & Vicari, 2009). The radial arm maze requires rats or children to remember which arms of a maze they have visited to retrieve a reward. This task recruits both working memory and spatial mapping abilities (for review see Goodman, Marsh, Peterson, & Packard, 2014). Work examining groups 2- to 5-year-olds, 6- to 12-year-olds, and adults has identified age-related improvements in performance and decreases in errors (Overman, Pate, Moore, & Peuster, 1996), with a rapid improvement around 5-years (Mandolesi et al., 2009). Taken together, these results demonstrate a consistent pattern of age-related improvement on VSWM tasks across childhood, but age-related differences during toddlerhood and the early preschool years have not been well-investigated.

**Sex-related Differences in Use of Spatial vs. Visual Cues**

Apart from age-related changes in VSWM, some researchers have also reported sex-related differences in the processing of spatial and visual cues. Boys are better at processing spatial cues than girls, and these sex-related differences emerge early in development and continue through middle childhood (Hood, 1995; Holmes, Nardi, Newcombe, & Weisberg, 2015; Moore & Johnson, 2008). In preschoolers, spatial performance is higher for boys than girls using
an established mental rotation task (Levine, Huttenlocher, Taylor, & Langrock, 1999). In later childhood, 5- to 7-year-old boys are better at perceiving and identifying changes in slope than girls (Holmes, et al., 2015).

In contrast, Mandolesi and colleagues (2009) showed that girls exhibited fewer errors and having a longer spatial span on the 8-arm radial maze. These authors concluded that girls show earlier acquisition of spatial exploration and navigation than boys (Mandolesi et al., 2009). Girls appear more capable than boys of using perceptual information such as color on spatial tasks. For example, Diamond et al. (1997) showed that 3.5- to 7-year-old girls performed better than boys on three tasks that involved integration of color cues with spatial cues. Specifically, girls performed better on the three pegs color tasks, the Corsi-Milner test, and the 6-item scrambled and stationary spatial search tasks. Similarly, Joh and Spivey (2014) tested 3-year-olds on the tubes task, a task that required participants to track the spatial path of a ball down a tube to predict where it would emerge from a crossed tube apparatus. They showed attenuation in sex differences when color cues were added to the tubes. Without these color cues, boys performed significantly better on the task than girls (Hood, 1995). Taken together, these studies reveal a male advantage for representing spatial information and a female advantage in integration of visual cues during spatial search.

The Present Study

To better understand developmental trajectories in the ability to update and manipulate perceptual and spatial cues, as well as any sex difference underlying performance early in development, we tested 2- to 4-year-old preschoolers on the Spin the Pots VSWM task. The Spin the Pots task is a multi-location search task that was derived from methodology used with non-human primates (Petrides & Milner, 1982) and developed by Hughes and Ensor (2005) to assess
working memory and inhibitory control during early childhood. Analogous to the scrambled versions of the *multiple boxes* task (Petrides & Milner, 1982), this non-verbal, internally ordered, and scrambled task requires children to retrieve 6 hidden objects from 8 distinctly colored cups. In *Spin the Pots*, color of the cup is an essential cue for children to encode and update across trials as location changes per trial. To succeed on the task, children must use VSWM to maintain and update their memory for cups where stickers have been retrieved. It addresses the limitations of other tasks because the non-verbal, internally ordered nature of the task is free from language constraints. Thus, the complexity is more age-appropriate relative to frequently used verbal WM tasks where toddlers may not understand or be unable to successfully hold instructions in WM.

*The Spin the* Pots task was developed and tested as part of a comprehensive executive functioning (EF) battery for 2 year-olds (Bernier, Carlson & Whipple, 2010; Hughes & Ensor, 2005). It is not clear from the current data how older children perform on the task. Hughes and Ensor (2009) reported that 4-year-old children in their sample performed at ceiling on the task and thus omitted them from analyses. A recent study of 120 children did not report age differences or find sex differences with 2- to 4-year-olds on *Spin the Pots* (Blakey, Visser, & Carroll, in press). There are differences across tasks regarding who completed the hiding phase; in some studies (Müller, Liebermann-Finestone, Carpendale, Hammond, & Bibok, 2012; Roman, Ensor, & Hughes, 2016), the children hid the stickers with the experimenter, while in others (Blakey et al., in press, Johansson et al, 2015) the experimenter hid the stickers, and in the original protocol it does not specify (Hughes & Ensor, 2005). Given the prior work on self-ordered WM tasks (see Pinto-Hamuy & Linck, 1965; Petrides & Milner, 1982, Diamond et al., 2007), our methods included a child-directed hiding phase prior to testing. This also allowed us
to assess whether or not children spontaneously used a hiding strategy to facilitate object retrieval.

In addition to differences in protocols, other limitations of previous studies using Spin the Pots include: reporting results by grouping EF measures into a composite score (see Hughes & Ensor, 2005; Bernier, Carlson, Deschênes, & Matte-Gagné, 2011), using the task at only one age (Bernier, Carlson, & Whipple, 2010) or not reporting certain descriptive preliminary analyses such as age (Blakey et al., in press) or sex (Johansson et al., 2015). To overcome these limitations we will examine these factors in a wide age-range using a more detailed coding system. Furthermore, a significant limitation of previous studies is that children’s performance for successfully completing the task has not been compared to random guessing. This limitation will be addressed by calculating the expected value of the number of trials to retrieve a prescribed number of stickers in the present study.

The aim of the present study is to characterize age-related changes and any associated sex difference in VSWM development from age 2 to 4 years using the Spin the Pots task (Hughes & Ensor, 2005). We hypothesized that with each additional year, children would become more efficient at solving the task (better working memory score and fewer perseverations). We hypothesized that if girls have an advantage integrating multiple cues, they would outperform boys on the task.

**Method**

**Participants**

Participants included 621 typically developing children (306 boys). Data were collected for 24 to 36 month-old children in their homes in Washington, DC, and 30- to 52-month-olds in the Smithsonian National Zoo and the Smithsonian National Museum of Natural History.
Portions of data collected with these participants have been published elsewhere (Subiaul, Zimmermann, Renner, Schilder, & Barr, 2015; Moser, Zimmermann, Dickerson, Grenell, Barr, Gerhardstein, 2015). Independent groups of children were tested at 2 years \((n = 334, M \text{ age} = 27 \text{ months 21 days}, SD = 3.53 \text{ months})\), 3 years \((n=201, M \text{ age} = 37 \text{ months 30 days}, SD = 2.97 \text{ months})\) and 4 years \((n=86, M \text{ age} = 51 \text{ months}, SD = 7.3 \text{ months})\). Participants were primarily Caucasian (68%) and from college-educated families \((M \text{ years of education} = 17.51, SD = 1.32)\). The remaining 29.5% of the sample included the following races: mixed (10.5%), African-American (4.8%), Asian (4.2%), Native American (0.2%), other (0.8%); the remaining participants \((n=11.6\%)\) did not report the race of the child. With regard to ethnicity, 11.9% of the sample was Hispanic. Approximately one third of children lived in homes where more than one language was routinely spoken (32.5%; bilingual homes). The mean rank of socioeconomic index (SEI; Nakao & Treas, 1994) was 79.96 \((SD = 11.42)\) based on 281 families (45.25%).

Productive vocabulary measured by the MacArthur Communicative Developmental Inventory (MCDI) from 341 children at 2 years \((\text{raw score } M = 66.95, SD = 23.80; n=254)\) and 3 years \((\text{raw score } M= 71.71, SD = 20.99, n=87)\). Fifty children were excluded from the analysis: 12 due to experimenter error, 3 for technical error with video, 11 for failure to interact with the experimental stimuli, 4 due to parental or sibling interference, and 20 for failure to complete the task.

**Apparatus and Stimuli**

The *Spin the Pots* (Hughes & Ensor, 2005; Bernier et al., 2010) apparatus is comprised of eight distinctly colored opaque cups, six attractive stickers, and a lazy Susan. All eight cups fit inside the lazy Susan in a circle with equal spacing between them. An opaque cover was used to cover the cups in between trials and had a handle on top of the cover in order to easily cover and
uncover the lazy Susan, see Figure 8. It was 15 cm tall, 35 cm in diameter, and 110 cm in circumference.

<table>
<thead>
<tr>
<th>Hiding Phase</th>
<th>Retrieval Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
</tbody>
</table>

*Figure 8. Hiding Phase and Retrieval Phase of Spin the Pots.*

**Self-report Measure**

After obtaining informed consent, the caregiver was asked to complete a general information questionnaire (assessing rank Socioeconomic Status (SES), parental education, and language exposure) as well as the MacArthur Communicative Development Inventory: Words and Sentences Short Form (MCDI) to measure children’s productive vocabulary (Fenson et al., 2000).

**Design and Procedure**

For the *Spin the Pots* task, the experimenter encouraged the child to place the six attractive stickers under six of the eight brightly colored cups, leaving two cups empty. After all stickers were hidden, the experimenter showed the child the two cups that did not have a sticker and said, “Look, no stickers under these cups!” The opaque cover was placed over all the cups
on the lazy Susan and the entire tray was spun 180 degrees. The experimenter uncovered the cups and instructed the child to find one of the stickers. If the child found a sticker, the experimenter praised the child, the sticker was set aside or given to the child’s caregiver, and the lid was replaced and the tray was spun 180 degrees again. After each trial, the tray was spun 180 degrees to counterbalance the position of the cups. If the child did not find a sticker, the experimenter gave appropriate feedback (e.g., “no sticker there, let’s try again”), the lid was replaced, and the tray was spun 180 degrees again. This task required the child to hold the location of the cups that did not have stickers in mind and to update this memory after each trial. The task ended when the child found all six stickers or reached sixteen trials. A subset of children did not find all six stickers within sixteen trials. Originally, researchers (Hughes & Ensor, 2005) gave children up to 16 trials to find all the stickers, but in the present study if the experimenter determined the child was able to continue with the task and had not yet retrieved all 6 stickers, they were tested for up to 35 trials ($M_{2years}=15.16$, $SD_{2years}=5.18$, $M_{3years}=14.43$, $SD_{3years}=4.86$, $M_{4years}=12.20$, $SD_{4years}=4.14$; $M_{boys}=15.18$, $SD_{boys}=5.14$, $M_{girls}=13.89$, $SD_{girls}=4.84$). For this reason, we used a new analytical approach to assess working memory and perseveration.

Coding

Task performance was videotaped for subsequent coding. For the Spin the Pots task, each child was given the following scores: adjusted WM, success rate, perseveration, alternate perseveration, non-perseveration search error, location, color strategy, and linear strategy.

Adjusted WM score. Using the approach developed by Hughes and Ensor (2005), the WM score was calculated as sixteen minus the number of errors made if the child found all six stickers or completed all sixteen trials. However, if the child did not find all six stickers or
complete all sixteen trials, their score was calculated based on the number of stickers found in the total number of trials completed. This was to ensure that a child’s score would not be inflated due to inability to complete the task.

**Success rate.** Given that the number of trials could vary across children, the success rate was calculated by dividing the number of stickers retrieved by the total number of trials.

**Perseveration.** This is the number of times the child chose a cup that was selected on the previous trial (whether it did or did not have a sticker on the first search). This allows us to quantify errors based on the feature of the cup. For example, selecting purple, then purple again across two trials would equate to one point on this measure.

**Alternate Perseveration.** This is the number of times the child chose a cup that was selected two trials ago (whether it did or did not have a sticker on the first search). This allows us to quantify errors based on the location of the cup as the apparatus rotated 180 degrees each trial. For example, selecting purple, then orange, then purple across three trials would equate to one point on this measure.

**Non-perseveration search errors.** This measure includes all other searches to an empty cup that was not selected one or two trials ago. It was calculated by taking the total number of trials minus 6 (for total stickers retrieved) minus the number of perseverations and alternate perseverations.

**Location.** The apparatus was divided in half on each test trial and the four cups closest to the child was considered “bottom” and the four cups further from them was called “top.” For each trial it was determined whether the child reached for a cup in the four nearest to them (bottom) or four on the other half of the apparatus furthest from them (top). The total number of
times a child searched in the top and bottom was computed. Only one selection was possible per trial.

**Color Strategy.** This measure quantifies hiding behavior of the child prior to the test phase. If the child matched at least four of the six stickers to their correct cup (i.e. green smile sticker under green cup) they received a point. Matching three or fewer stickers was not defined as color strategy use.

**Linear Strategy.** This measure quantifies hiding behavior of the child prior to the test phase. If the child hid at least four of the six stickers in a linear fashion (without skipping cups) around the circumference of the apparatus they received a point. Linear hiding of 3 or fewer stickers was not defined as location strategy use.

**Reliability.** A second independent coder scored 14% of the videos to determine reliability of the ratings; coders recorded the color of the cup selected per trial and whether or not a sticker was retrieved. Inter-rater reliability on success rate was calculated (kappa = 0.92).

**Results**

Preliminary analyses including bilingual status and parental education were conducted for each of the dependent measures (adjusted working memory score, success rate, perseveration, and alternate perseveration). No significant ($p < .05$) effects or interactions involving bilingual status or parental education emerged in the preliminary analyses, and data were collapsed across these variables in further analyses.

**Expected value**

A limitation of the current task is that previous studies have not compared performance to the statistical expected value of the number of trials to retrieve a prescribed number of stickers. We calculated the expected value of the probability distribution of retrieving a prescribed
number of stickers in 200 trials using six million Monte Carlo simulations. We considered the cases of retrieving five and six stickers. Children’s performance was compared to the expected values for each of these cases. This calculation allowed us to assess whether children, in the final sample, were in fact recruiting VSWM to complete the task. The expected value to retrieve all six stickers is 18.35 trials, \( SD = 6.69 \) and to retrieve five stickers 11.6 trials, \( SD = 4.44 \).

To test whether performance differed by age, sex, and number of stickers retrieved when the expected value was taken into account, a centered mean was computed for total trials by subtracting trials from 18.35 for those who retrieved 6 stickers and subtracting trials from 11.6 for those who only retrieved 5 stickers.

A 3(age: 2, 3, 4 years) x 2(sex: male, female) x 2(stickers: 5, 6) ANOVA across the centered mean revealed a main effect of age, \( F(2,586) = 4.69, p < .05, \eta^2_p = .016 \), a main effect of sex, \( F(2,586) = 4.58, p = .01, \eta^2_p = .008 \), a main effect of stickers, \( F(2,586) = 436.78, p < .001, \eta^2_p = .43 \). There were no significant two way interactions between sex and age \( F(2,586) = .33, p = ns \), sex and stickers \( F(2,586) = .72, p = ns \), or age and stickers \( F(2,586) = .49, p = ns \). These findings indicate there is a total trials (corrected for the expected value) and age as well as the number of stickers. It is not clear from ANOVA which ages and which sticker numbers were significantly different from the expected value of trials from the Monte Carlo simulation. Therefore, follow-up t-tests are needed to determine which groups performed better or worse than the expected value in terms of total trials to retrieve the stickers.

Under the null hypothesis, we would expect that each age group would be performing at the expected value. However, we are able to reject the null hypothesis \( (p < .001) \) for all age groups. Namely, children \( (n = 446) \) who retrieved six stickers performed significantly better than the expected value at 2, 3, and 4 years \( (t(230) = -10.29, p < .001, t(144) = -9.45, p < .001, t(69) = - \)
8.87, \( p < .001 \), respectively). In other words, they retrieved six stickers in fewer trials than the expected value of 18.35 trials. This suggests that all age groups who retrieved six stickers were significantly different from the expected value of 18.35 and thus we interpret these findings that children were not simply guessing randomly, but attempting to keep track of and update information across trials. The percentage of children who found six stickers relative to the entire sample by age is as follows: 2-year-olds: 69.16\% (231/334), 3-year-olds 72.14\% (145/201), and 4-year-olds 81.4\% (70/86). Furthermore, t-tests were also calculated separately for males and females at each age. The pattern of results was the same at each age and for each sex of the child. However, this was not true for children who only retrieved five stickers. Across age groups and both sexes, children who retrieved 5 stickers did not complete the task in fewer trials than the expected value of 11.36 trials; all t’s not significant. Given that children who only found five stickers did not perform better than the expected value, their data were not analyzed further.

Results from the simulations were also used to determine the expected number of trials until the first error. The expected number of trials until the first error is 1.33 (\( SD = .67 \)). All age groups had their first error significantly later than the expected first error, 2-year-olds: \( t(331) = 36.4, p < .001 \), 3-year-olds: \( t(196) = 29.67, p < .001 \), and 4-year-olds: \( t(80) = 25.79, p < .001 \) (see Table 6). Again, given that the first error occurred significantly later than expected, we interpret this finding to indicate that children were not simply guessing at random.

The mean scores and standard deviations for the Spin the Pots task as a function of age in years and sex of the child are reported in Table 1 for those children who retrieved 6 stickers (\( n = 446 \)).
Table 6. Mean time to complete task, success rate, adjusted working memory score, total trials, and (SDs) as a function of age and sex of the child for those who retrieved six stickers.

<table>
<thead>
<tr>
<th>Sex</th>
<th>Age</th>
<th>Time to complete (s)</th>
<th>Success Rate</th>
<th>Adjusted WM score</th>
<th>Total trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>2</td>
<td>161.79 (68.52)</td>
<td>.44 (.14)</td>
<td>8.08 (3.20)</td>
<td>14.83 (4.75)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>129.67 (59.17)</td>
<td>.50 (.19)</td>
<td>8.92 (3.63)</td>
<td>13.83 (4.97)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>106.69 (34.76)</td>
<td>.53 (.16)</td>
<td>9.80 (3.19)</td>
<td>12.20 (3.19)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>143.49 (64.93)</td>
<td>.47 (.16)</td>
<td>8.60 (3.37)</td>
<td>14.12 (4.68)</td>
</tr>
<tr>
<td>Female</td>
<td>2</td>
<td>146.91 (65.74)</td>
<td>.50 (.17)</td>
<td>9.26 (3.53)</td>
<td>13.32 (4.67)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>122.79 (45.22)</td>
<td>.51 (.16)</td>
<td>9.46 (3.35)</td>
<td>12.87 (4.08)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>90.23 (34.03)</td>
<td>.66 (.20)</td>
<td>11.97 (3.61)</td>
<td>10.20 (4.11)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>136.37 (58.40)</td>
<td>.53 (.18)</td>
<td>9.74 (3.59)</td>
<td>12.68 (4.50)</td>
</tr>
</tbody>
</table>

Total time to complete task

A 3(age: 2, 3, 4 years) x 2(sex: male, female) ANOVA across total time to complete task revealed a main effect of age, $F(2,445) = 27.76, p < .001, \eta^2_p = .79$, large effect size, with older children completing the task at a faster rate. There was also a main effect of sex of the child $F(1, 445) = 4.21, p < .05, \eta^2_p = .11$, medium effect size, with girls completing the task faster, but no significant age x sex interaction $F(2,445) = .26, ns$.

Total Trials to complete task

A 3(age: 2, 3, 4 years) x 2(sex: male, female) ANOVA across total trials to complete task revealed a main effect of age, $F(2,445) = 11.98, p < .001, \eta^2_p = .04$, small effect size, with older
children completing the task at a faster rate. There was also a main effect of sex of the child $F(1, 445) = 6.78, p < .01, \eta^2_p = .01$, small effect size, with girls completing the task faster, but no significant age x sex interaction $F(2, 445) = .09, ns$.

**Success rate**

A 3(age: 2, 3, 4 years) x 2(sex: male, female) ANOVA revealed a main effect of age, $F(2, 445) = 14.04, p < .001, \eta^2_p = .06$, medium effect size, a main effect of sex of the child $F(1, 445) = 15.01, p < .001, \eta^2_p = .03$, small effect size, with females outperforming males, and a trend toward a significant age x sex interaction $F(2, 445) = 2.67, p = .07$ (see Figure 9). A post-hoc Student Newman-Keuls (SNK, $p < .05$) analysis across all groups indicated that the 4-year-olds ($M = .59, SD = .19$) significantly exceeded the performance of the 2- and 3-year-olds ($M_2 = .47, SD_2 = .16; M_3 = .50, SD_3 = .18$). There was no difference between 2- and 3-year-olds, see Figure 2.

![Success rate working memory scores (±1SE) as a function of age and sex for participants who retrieved 6 stickers.](image)

*Figure 9.* Success rate working memory scores (±1SE) as a function of age and sex for participants who retrieved 6 stickers.
Table 7. Perseveration, alternate perseveration, non-perseveration errors, and first error (SDs) as a function of age and sex of the child for those who retrieved 6 stickers.

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Perseveration</th>
<th>Alternate Perseveration</th>
<th>Non-perseveration Errors</th>
<th>First error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>2</td>
<td>.74 (1.19)</td>
<td>2.75 (2.69)</td>
<td>5.35 (3.46)</td>
<td>2.62 (1.33)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.70 (1.01)</td>
<td>2.56 (2.32)</td>
<td>4.57 (3.27)</td>
<td>2.47 (1.23)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.77 (.91)</td>
<td>1.31 (1.21)</td>
<td>4.11 (2.93)</td>
<td>3.14 (1.46)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>.73 (1.09)</td>
<td>2.46 (2.45)</td>
<td>4.92 (3.34)</td>
<td>2.66 (1.33)</td>
</tr>
<tr>
<td>Female</td>
<td>2</td>
<td>.62 (.82)</td>
<td>2.28 (2.18)</td>
<td>4.42 (3.26)</td>
<td>2.76 (1.51)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.68 (.95)</td>
<td>1.94 (1.85)</td>
<td>4.24 (2.80)</td>
<td>3.01 (1.30)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.23 (.55)</td>
<td>1.20 (2.07)</td>
<td>2.77 (2.39)</td>
<td>3.55 (1.15)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>.58 (.85)</td>
<td>2.00 (2.08)</td>
<td>4.10 (3.03)</td>
<td>2.96 (1.41)</td>
</tr>
</tbody>
</table>

Perseveration

A 3(age: 2, 3, 4 years) x 2(sex: male, female) ANOVA revealed no main effect of age, $F(2, 445) = 1.06, ns$, but did show a main effect of sex of the child $F(1, 445) = 4.79, p = .03, \eta^2_p = .01$, small effect size, with fewer perseveration errors for females, and no significant interaction $F(2, 445) = 1.81, ns$.

Alternate Perseveration

A 3(age: 2, 3, 4 years) x 2(sex: male, female) ANOVA revealed a main effect of age, $F(2, 445) = 8.60, p < 0.001, \eta^2_p = .038$, small effect size, no main effect of sex of the child $F(1, 445) = 2.84, p = ns$, and no significant interaction $F(2, 445) = .30, ns$. A post-hoc Student Newman-Keuls (SNK, $p < 0.05$) analysis across all groups indicated that the 4-year-olds ($M =$
1.26, \( SD = 1.68 \) had significantly fewer alternate perseveration errors than the 2 and 3-year-olds \((M_2 = 2.52, SD_2 = 2.46; M_3 = 2.21, SD_3 = 2.08)\), who did not differ from each other.

Figure 10. Perseveration (left) and Alternate Perseveration scores (right) (±1SE) as a function of age and sex for participants who retrieved 6 stickers.

Non-perseveration errors

A 3(age: 2, 3, 4 years) x 2(sex: male, female) ANOVA revealed a main effect of age, \( F(2, 445) = 5.69, p < .01, \eta^2_p = .025 \), small effect size, a main effect of sex of the child, \( F(1, 445) = 6.64, p < .01, \eta^2_p = .015 \), small effect size, with fewer non-perseveration errors for females, and no significant interaction \( F(2, 445) = .71, ns \). A post-hoc Student Newman-Keuls (SNK, \( p < 0.05 \)) analysis across all groups indicated that the 4-year-olds \((M = 3.44, SD = 2.74\) had significantly fewer non-perseveration errors than the 2 and 3-year-olds \((M_2 = 4.89, SD_2 = 3.39; M_3 = 4.39, SD_3 = 3.01\), who did not differ from each other.
Location

There was a significant positive correlation between choosing cups located in the bottom half of the apparatus (closest to the child) on the test trials and higher rates of perseveration, \( r(446) = .24, p < .001 \) and alternate perseveration, \( r(446) = .73, p < .001 \). Older children (\( r(444) = -.27, p < .01 \)) and girls (\( r(444) = -.13, p < .01 \)) were significantly less likely to select cups from the bottom.

Vocabulary

To test whether productive vocabulary, age, and sex, predicted working memory success rate scores a 2 x 2 ANOVA with MCDI as a covariate was conducted. A subset of the sample \( n=341 \) 2- to 3-year-olds had parents who completed the MCDI, a measure of productive vocabulary. A 2(age: 2, 3 years) x 2(sex: male, female) ANOVA with MCDI as a covariate on success rate did not yield any significant main effects or interactions for MCDI \( F(1,340) = .67, p=ns \), sex of the child, \( F(1, 340) = .02, p=ns \), age \( F(1,340) =3.70, p=ns \), no significant interactions between sex and MCDI \( F(1,340) =.20, p=ns \), age and MCDI, \( F(1,340) =3.76, p=ns \), or sex and age \( F(1,340) =2.11, p=ns \). These results indicate that individual differences in productive vocabulary for 2- and 3-year-olds did not predict working memory scores and they did not differ by sex.

Strategy

Color. 91 children (20.4%) used a color matching strategy with 4 or more stickers. Girls and boys were equally likely to use a color strategy (\( n_{girls}=45; n_{boys}=46 \)) and 3- and 4-year-old children had more frequent color strategy use (\( n_2=35 \) (10.5%), \( n_3=41 \) (20.4%) \( n_4=15 \) (17.4%). A chi-square test of independence was performed to examine the relationship between age and color strategy. The relationship between these variables was significant, \( \chi^2 (2, N = 433) = 9.95 \),
\( p < .01, \phi = .15 \), small effect size. There was a significant positive correlation between children’s color matching stickers to the cups and success rate, \( r(433) = .13, p < 0.01 \).

Multiple regression analysis was used to test if color strategy, age, and sex predicted children’s success rate. The results of the regression indicated that predictors explained 9% of the variance (\( R^2 = .09, F(3,429) = 13.595, p < .001 \)). It was found that color strategy significantly predicted success rate (\( \beta = .11, p < 0.05 \)), as did age (\( \beta = .21, p < 0.001 \)) and sex (\( \beta = .16, p < 0.001 \)). Additional models were run that included interaction terms but in none of these models were the interaction terms between color strategy and sex or color strategy and age significant.

There was no significant correlation between perseveration and children’s color matching \( r(433) = -.05, ns \) or between alternate perseveration and color matching \( r(433) = -.05, ns \).

**Linear.** 139 children (32.1%) used a linear hiding strategy with 4 or more stickers. Girls and boys were equally likely to use a linear strategy (\( n_{girls} = 71; n_{boys} = 68 \)) and 3- and 4-year-old children had more frequent linear strategy use (\( n_2 = 55 (16.5\%), n_3 = 54 (26.9\%) n_4 = 30 \) (34.9%). A chi-square test of independence was performed to examine the relationship between age and linear strategy. The relationship between these variables was significant \( \chi^2 (2, N = 437) = 13.33, p < .01 \). There was no significant correlation between a linear hiding strategy and success rate \( r(437) = -.05, ns \). There was also no significant correlation between perseveration and linear hiding strategy, \( r(437) = .03, ns \) or alternate perseveration and linear hiding strategy, \( r(437) = .01, ns \).
Table 8. Correlation Table

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Color</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>strategy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Linear</td>
<td>-</td>
<td>-.28**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>strategy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Success</td>
<td>.13**</td>
<td>-</td>
<td>-.05</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Perseveration</td>
<td>-</td>
<td>.03</td>
<td>-.41**</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5. Alternate</td>
<td>-</td>
<td>.01</td>
<td>-.57**</td>
<td>.04</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Perseveration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Age</td>
<td>.09*</td>
<td>.17**</td>
<td>.23**</td>
<td>-.05</td>
<td>-.18**</td>
<td>-</td>
</tr>
<tr>
<td>7. Sex</td>
<td>-.01</td>
<td>.001</td>
<td>.17**</td>
<td>-.08</td>
<td>-.10*</td>
<td>.03</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01

Discussion

The present study increased our understanding of typical age and sex-related changes in VSWM development across the early childhood period. The present study measured performance across an extended age range during early childhood allowing for additional trials to complete the task, and compared overall performance of children successfully completing the task to random guessing in order to improve interpretation of task performance. Consistent with our hypotheses, we found both age-related and sex-related differences in VSWM. The study established that by age 4, children are performing significantly better than 2- and 3-year-olds on a self-directed hiding version of the Spin the Pots VSWM task. Furthermore, results indicated
that, across different measures, girls outperformed boys on this task. While this difference was significant, the effect size was small for success rate and perseveration.

The significant age-related findings and large effect sizes from the present study differ from previous *Spin the Pots* studies in this age range (e.g. Blakey et al., 2015) that did not report age differences or find sex differences with 2- to 4-year-olds, as well as others who report performance was at ceiling at 4 years (e.g. Hughes & Ensor, 2009). One reason our results may differ could be due to methodological differences during the hiding phase. In the present study children participated in a self-ordered hiding task rather than an experimenter directed or assisted hiding phase. While this self-ordered hiding might make retrieval easier for some children, additional working memory demands may have been present for those children who took longer to hide their stickers or who did not try to remember the color of the cups.

The observed age-related changes on the *Spin the Pots* task are consistent with studies examining neural correlates of the developing visual WM system. Specifically, in a functional near-infrared spectroscopy (fNIRS) study, Buss, Fox, Boas, and Spencer (2014) showed that increasing behavioral performance by 4-year-olds was associated with more robust parietal responses compared to 3-year-olds on a visual WM change detection task. Future work is needed to understand how behavioral age-related changes in VSWM are related to regions of activation and connectivity patterns in the brain.

In our analyses of inhibitory control, as measured by perseveration, we found girls perseverated significantly less than boys, although the effect size was small. This sex difference in the ability to inhibit a prepotent motor response or focus on relevant stimuli in the presence of distractors is consistent with the literature in preschool through early childhood. Previous research documented on multiple object search working memory tasks (Diamond et al., 1997;
Mandolesi et al., 2009) has identified sex differences in inhibitory control. A meta-analysis of children 3 months to 13 years has also documented a moderate sex difference for inhibitory control, suggesting that girls have a better ability to control responses or behaviors than boys (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006). These findings are consistent with the present study where we also document a lower rate of perseveration in females, even though the rate of perseverations was low overall. Overall our results suggest a lower overall error rate for girls. Specifically, our results indicated that girls completed the task in significantly fewer trials, took significantly less time, had significantly fewer perseverations, and had significantly fewer non-perseveration errors. There may be a female advantage for keeping track of visual features such as colors, shapes, or other features of stimuli which manifests in fewer overall errors due to either a hiding strategy, an efficient search strategy (where kids try to remember the color of the cups they have looked under previously), or both.

Furthermore, we speculate that the advantage for girls on these tasks and the Spin the \textit{Pots} task was due to better integration of color cues during the spatial search tasks, which resulted in fewer search errors. Due to the small effect size this sex difference must be interpreted with caution, especially due to the large sample size. An open research question remains: to test this speculation to measure to what extent young children rely on spatial or color information or integrating both. It is possible to complete the task simply by maintaining and updating color information, so children do not need to keep track of spatial information to succeed on the task, although spatial location does vary across trials. We explored this possibility in a post-hoc analysis by examining strategy use during the hiding phase of the task. We found that many of the children did not take a strategic approach to hiding the stickers and randomly placed the stickers under the pots. Some children, however, behaved in more deliberate ways.
during this phase, some placing the stickers in a linear fashion and others matching the color of
the sticker to the color of the pots. The color matching strategy was associated with a higher
success rate on the task, as might be expected. While only a small subset of the entire sample
(~20%) used this strategy, it significantly predicted a better success rate on the VSWM task but
was not associated with perseveration rates. This suggests that several different factors may be
involved in predicting WM performance.

The findings with regard to MCDI are not surprising. While in Chapter 2 we see the
importance of MCDI for HTP, this is not the case in 2- to 3-year-olds in predicting success on
the task. This may be due to rapid changes in language acquisition in the first two years of life,
with less significant changes after the age of 2. Another possible interpretation is the complex
nature of this task is such that a large portion of the sample at 2 (30.84%) and 3 (27.86%) did not
find all 6 stickers. This also may be due to the fact that MCDI at 3 years was not significantly
different from 2 years. We can conclude that the sex differences on this task are not due to
differences in productive vocabulary. This is in contrast to experiment 1 study 1 (Chapter 2)
where MCDI was a better predictor of Hide the Pots scores than age at 18 and 24 months. These
findings suggest that vocabulary differences early may underlie performance on simple working
memory tasks but as kids age and working memory tasks become more complex these
differences may no longer be associated. It is an open question whether this is due to the
complex nature of the task, as evidenced by poorer performance in 2- and 3-year-olds relative to
4-year-olds, or due to less variation in productive vocabulary in toddlerhood relative to later in
infancy. One limitation is that MCDI is not a valid assessment for 4-year-olds. Perhaps if another
vocabulary assessment were used at this age we may see it predicts performance for those
individuals with a higher success rate.
Limitations and future directions

One limitation of the task is its complexity. While Spin the Pots was initially designed and tested with 2-year-olds, our results suggest that performance is poor at this age. This may be due to the complex nature of the task. One reason is that the spatial location changes each trial as the apparatus is rotated 180 degrees (similar to scrambled boxes). A second reason is that two cups remain empty, meaning that children have to keep track of and update information about the two empty cups in addition to the new empty cups from which they retrieve stickers. Scrambled boxes may be more challenging because there are documented age-related changes in children’s ability to both use landmarks to track objects, and to track objects through spatial rotation to infer locations of hidden objects (Okamoto-Barth & Call, 2008). Specifically, in a two location search task, 3-year-olds performed well on visible displacements and invisible rotation if a marker was on top of the target cup, but poorly with invisible displacement in the absence of landmark cues. Inferring rotations was achieved later in development, with 5-year-olds tracking 180 degree rotations independent of landmark cues. This partially accounts for why tracking 8 locations even with highly distinctive cups in the Spin the Pots task is challenging for 2- to 4 year-olds.

Additional follow-up studies are needed to better understand factors that impact VSWM performance during the preschool years. Factors to further investigate include the role of delay between hiding and retrieval event and the spatial location of objects by comparing stationary pots (location fixed throughout the task) versus Spin the Pots (location of cups relative to the child varies on each trial). Additional factors to examine include self-ordered versus experimenter-ordered hiding prior to the test phase and cognitive load by varying the number of locations and/or the number of hidden objects to retrieve. Future research could systematically
assign children to different hiding strategy conditions with experimenters modeling a strategy such as color matching to examine how this information influences performance as a function of age and sex of the child. These findings would be relevant to early educators as potential ways to improve performance by children who are having difficulty keeping track of the information. Lastly, work looking at predicting academic outcomes has used verbal WM tasks but future longitudinal studies should incorporate non-verbal WM measures.

This systematic approach has been useful in prior studies which have incorporated modifications to the task parameters to make it simpler, more complex, or otherwise modified for different ages. An easier version of Spin the Pots comes from a study that incorporated a variant of Spin the Pots with 2-item, 3-item, and 4-item trials (adapted from Diamond et al., 2007) but only measured children at 24 months (Hammond, Müller, Carpendale, Bibok, & Liebermann-Finestone, 2012). A more complex version of Spin the Pots incorporated 11 cups and 9 stickers and was used with 4.5- to 5.6-year-olds (Batchelor, Inglis, & Gilmore, 2015). Another study incorporated a variant of Spin the Pots where the number of boxes and trials were age dependent: 8 boxes, 12 spins (2.5 yrs), 9 boxes, 14 spins (3 yrs), 10 boxes, 16 spins (4 yrs), 11 boxes, 18 spins (4 yrs), but only reported results within an EF composite (Hostinar, Stellem, Schaefer, Carlson, & Gunnar, 2012). Yet another followed a similar age dependent approach of 8 boxes for ages 2.5–3.5, 10 boxes for ages 3.5–4.5, and 12 boxes for ages 4.5–5.5 with a maximum number of spins determined to be 12 spins for ages 2.5–3.5; 16 spins for ages 3.5–4.5; 20 spins for ages 4.5–5.5 (Spann & Gagne, 2015). They did not report age differences and did not find sex differences in WM.
Implications

The present findings add to a small but growing body of literature on developmental and sex-related changes in VSWM during early childhood. The development of visuospatial WM is crucial for acquiring skills in complex tasks such as mathematics and problem solving that require substantial information tracking and updating, which ultimately impact academic success. WM and spatial skills are critical for early education and are precursors to success in Science, Technology, Engineering, and Math (STEM) disciplines. There are significant differences in the home environment in activities that enhance spatial transformation skills. For example, parents of 2- to 4- year-olds are both more engaged and use more spatial language with boys than they do with girls, and expose boys to more difficult puzzles than girls (see Levine, Ratliff, Huttenlocher, & Cannon, 2012). More frequent puzzle play was associated with better spatial transformation for both boys, and higher puzzle quality was associated with better spatial transformation for girls. It is important to integrate findings from spatial tasks that typically show male advantages for representing spatial information (e.g. Levine et al., 1999) with others that show a female advantage in integration of cues during spatial search (e.g. Diamond, 1997; Mandolesi et al., 2009). The reported age differences also have important implications for early education. Specifically, educators should consider different strategies to enhance spatial learning for younger children who have greater difficulty on these tasks. Perhaps the addition of color cues during tasks that have spatial elements such as number lines, sorting objects, pattern detection, mental rotation, puzzles, and the translation of geometric shapes may facilitate performance in boys and girls by taking advantage of their ability to integrate spatial and color information successfully at an early age.
In conclusion, the current research tested 2-, 3-, and 4-year-old children’s memory for hidden objects based on their visual and spatial perceptual cues in the short-term. Importantly, our study is the first to incorporate a statistical approach to compare performance to an expected value for the total trials to complete Spin the Pots. Based on this approach, we determined that all age groups who found 6 stickers were significantly different from the expected value of 18.35 trials and thus children were not simply guessing randomly but attempting to keep track of and update information across trials. Age-related and sex-related differences were reported, with both girls and older children completing the task more successfully and with fewer perseveration errors. However, this was not true for children who only retrieved 5 stickers. Spontaneously employing a color matching sticker hiding strategy use was associated with better a success rate on the task.
CHAPTER IV: A NEW TOUCHSCREEN VISUOSPATIAL WORKING MEMORY TASK AND ITS’ RELATIONSHIP TO INHIBITORY CONTROL AND MOTOR-SPATIAL IMITATION IN PRESCHOOLERS

Working memory involves the ability to retain and manipulate information for a short time. A subcomponent of working memory, Visuospatial Working Memory (VSWM) involves an ability to maintain and update non-verbal, spatial information on the scale of seconds. Despite the creation of numerous working memory tasks for children (see Chapter 1, Table 1), it has been challenging to document age-related changes in VSWM in preschoolers. Part of the challenge can be attributed to creating a valid, age-appropriate measure to assess VSWM.

There are several widespread assessment challenges. First, many working memory tasks designed for preschoolers, such as forward and backward digit or word span and the NIH Toolbox List Sorting Working Memory Test, rely heavily on verbal skills to succeed. Specifically, the LSWM test requires complex verbal recall and performance is notably poor in 3- to 5-year-olds (Riggins, 2013). Furthermore, object-based VSWM tasks for preschoolers such as Spin the Pots (Hughes & Ensor, 2005) or radial arm maze (Mandolesi, Petrosini, Menghini, Addona, & Vicari, 2009) are time consuming and can be difficult to explain to children. In addition, they are not easily adaptable to neuroimaging measurement, such as near-infrared spectroscopy (NIRS) due to the large amount of motion required for the retrieval of hidden objects.

Touchscreen as an assessment solution. To address the problems of verbal report or object-based tasks described earlier, some researchers have recently attempted to measure cognitive development via touchscreens rather than object based tasks. One recent study documented that touchscreens are a viable method to use to collect data from children between 1 and 4 years (Frank, Sugarman, Horowitz, Lewis, & Yurovsky, 2016). Another recent study
documented the effects of an EF training program for preschoolers using 4 touchscreen tasks (Blakey & Carroll, 2015). Taken together, this suggests that using the touchscreen to assess emerging EF in the preschool age-range is both an ecologically valid and viable option.

**CANTAB task.** Although not specifically designed for children, researchers have recently adopted the Cambridge Neuropsychological Test Automated Battery Spatial Working Memory Task (CANTAB) (Sahakian & Owen, 1992) to examine changes in working memory in children as young as 4 years of age (Luciana & Nelson, 1998). Specifically, during the CANTAB Spatial Working Memory (SWM) task children are presented with a series of 40 trials of increasing task complexity (from 2 to 8 items) presented on a touchscreen. The child’s task is to find hidden tokens behind filled squares on the screen while ignoring squares where they previously retrieved a token (Luciana et al., 1998). Researchers calculated a span score and a perseveration score (Luciana & Nelson, 2002). Results indicated that VSWM has a long developmental trajectory and performance did not reach adult-like levels until 12 years of age (Luciana et al., 2002). Although it is a validated and normed assessment tool, it was difficult to test 4-year-olds because of the complexity of the task demands and limited touchscreen sensitivity. To address these limitations, an age-appropriate modification to the CANTAB SWM was created as a touchscreen app to examine age-related changes in VSWM in preschoolers (3- to 5-years).

*Figure 11. CANTAB Spatial Working Memory Task*
**Relationship between Working Memory, Inhibitory Control, and Motor-spatial Imitation**

Early childhood is thought to be a time period during which executive functions (EF), which include working memory, inhibitory control, and shifting rapidly develop, with simple abilities becoming coordinated thus allowing for more complex abilities (Garon, Bryson, & Smith, 2008; Garon, Smith, & Bryson, 2014). This hierarchical model of EF development proposed by Garon and colleagues (Garon et al., 2008) states that attention serves as a foundation. Furthermore, simple skills such as maintaining information or delaying response begin to emerge in the first three years of life. Then, as these components are integrated into a complex ability such as working memory or inhibition they begin to become more mature in the preschool years. The present study sought to understand how executive functions develop across early childhood using the hierarchical model that would support the prediction of age-related changes as cognitive load (or task complexity) increases. This study will lay the foundations for understanding basic developmental changes and task parameters that relate to performance. This will enable future work regarding the complex relationship between EF components of working memory and inhibitory control. The relationship between working memory and inhibitory control is still not well understood as there is disagreement if inhibition is separate from WM, a product of exercising WM, or WM and inhibition depend on the same limited capacity system and thus are linearly related (see Diamond, 2013 for review).

Researchers have reported that components of working memory are related to 1 to 2 year-olds (Bernier, Carlson, & Whipple, 2010), as well as 2 and 3 year-olds (Hughes & Ensor, 2005) inhibitory control. Blakey and Carroll (2015) found that working memory training on the Six Boxes object-retrieval task was associated with training gains from the Flanker inhibitory control task. Tsujimoto and colleagues (Tsujimoto, Kuwajima, & Sawaguichi, 2007) also found that
VSWM was significantly correlated with performance on their inhibitory control go/no-go task in 5- to 6-year-olds, but not 8- to 9-year-olds. Working memory load has also been examined within the context of a cognitive imitation touchscreen task, where children view an experimenter demonstrate a sequence by touching unique pictures in order and across trials the spatial locations vary but pictures remain the same (Subiaul & Schilder, 2014). Specifically, results suggested WM constrains imitation as when demands were low (2-pictures) preschoolers copied order significantly more than when WM demands were high (3-pictures) on this touchscreen task.

While associations between working memory and motor-spatial imitation as well as working memory and inhibitory control have been documented in preschool, an association between the three domains of VSWM, inhibitory control, and motor-spatial imitation has not been investigated in the preschool years. To answer this question and address limitations associated with object-based tasks and other tasks that rely heavily on verbal report in the preschool years, the three tasks used will all be on a touchscreen. In addition to assessing behavioral performance on EF tasks, it is important to also assess parents reported observations of their children. This is because parent reports to EF questionnaires capture information regarding their observations of the child, making them ecologically valid. Furthermore, EF questionnaires (e.g. BRIEF-P) have been correlated with preschoolers EF behavioral performance (Garon, Piccinin, Smith, 2015) in typically developing 3- to 5-year-olds.
Present Study

The aim of the present study is to develop and assess performance on new VSWM task, *Find the Stars*. This task was modified based on the CANTAB SWM (Luciana & Nelson, 1998). While previous studies have reported poor performance in preschool (Luciana & Nelson, 1998; Luciana & Nelson, 2002) the task was adapted for the present study. To address previous limitations when testing young children, this new version has created modifications to stimuli, reduction of cognitive load, increased transparency of the task, and improved technical functionality. One main benefit of using this task over other working memory tasks is that it relies on non-verbal responses. These modifications will allow age-related changes in preschoolers to be better assessed and ultimately compared to two established tasks to examine the relationship between VSWM, inhibitory control, and motor-spatial imitation.

Inhibitory control will be measured using the Flanker task (Eriksen & Eriksen, 1974). More recently, it has been adapted for use with children as part of the NIH Toolbox cognition battery to measure executive function and attention (Zelazo, Anderson, Richler, Wallner-Allen, Beaumont, & Weintraub, 2013). Since this task was not developed for the touchscreen Courage and her colleagues (Frizzell & Courage, 2015) recently developed a modified version to make the task appropriate for use with preschoolers. This task requires children to press an arrow that corresponds with the direction the center fish is swimming, while the surrounding fish may be swimming the same or different direction. In their initial results, improvement with age was evident between 3 and 5 years (Frizzell & Courage, 2015).

Motor-spatial imitation will be measured using a well-established imitation paradigm (Subiaul, Anderson, Brandt, & Elkins, 2012) developed and tested with preschoolers in several studies. This task has a demonstration component, where children view an experimenter touching
identical pictures in a predetermined sequence, followed by a test phase. Numerous studies have shown that in early childhood there are age-related changes in motor-spatial imitation performance (e.g. Subiaul, Patterson, Renner, Schilder, & Barr, 2015; Subiaul, Zimmermann, Renner, Schilder, & Barr, 2015). Of particular interest is the spatial nature of this task as it requires participants to keep track of and update spatial information presented by an experimenter much like object-based working memory tasks do. In recent work it has been linked to working memory (Subiaul & Schilder, 2014). It will allow us to examine the relationship between spatial imitation and visuospatial working memory.

The BRIEF-P is a parental screening tool developed by clinicians to help determine whether children are experiencing difficulties in executive function. Parent reports on the BRIEF-P have been linked to clinically diagnosed specific learning disability, attentional deficit, memory impairment, or other diagnoses (BRIEF-P; Gioia, Espy, & Isquith, 2003). Parents report on their child’s behavior in a series of questionnaire items based on real world experience. We sought to examine whether parent reported EF would be associated with any of the established touchscreen tasks or the newly developed Find the Stars task.

It is predicted that 3 to 5 year-olds will be able to succeed on the Find the Stars task because of task modifications. It is also predicted that performance on Find the Stars will improve with age and vary based on cognitive load. Lastly, it is predicted that better performance on Find the Stars will be associated with better inhibitory control and better motor-spatial imitation on two established touchscreen tasks. It is predicted that BRIEF-P parent report measures will be associated with the EF tasks, specifically Find the Stars and the working memory subscale as well as Flanker Fish and the inhibit subscale.
Method

Participants

80 children between 3- and 5-years were recruited and tested in the National Building Museum in Washington, DC. Participants were primarily Caucasian (55%) and from graduate school-educated families ($M$ years of education = 17.05, $SD$ = 1.42). The remaining 19.8% of the sample included the following races: mixed (12.5%), African-American (3.8%), Asian (8.8%), Native American (1.3%), other (1.2%), and not reported (16.3%). With regard to ethnicity, 5% of the sample was Latino. Children exposed to a second language and were reported to have moderate to fluent proficiency were a small portion of the sample (14.1%). Touchscreen exposure ranged from never (11.5%), once a week (19.2%), 2-3 times per week (38.5%), up to 30 minutes per day (7.7%), 30-60 minutes per day (12.8%), more than 1 hour per day (10.3%), to no response (2.5%). Additionally, TV exposure per day ranged from never (15.4%), 30-60 minutes (48.8%), 1-2 hours (25%), 2-3 hours (7.5%), more than 3 hours (1.4%), to no response (2.5%). Behavior Rating Inventor of Executive Function- Preschool data was collected from 65 parents or caretakers (81.25%).

Apparatus and Stimuli

The Find the Stars task and Motor-spatial task were presented to participants on a Surface Pro 2 tablet (10.6 inch multitouch-screen display) and the Flanker Fish task was presented on an iPad Air (9.7-inch Retina display).

Design and Procedure

The battery of tasks was completed in a single session lasting 10 to 15 minutes per child. The order of the tasks was counterbalanced across participants. The protocols for the three touchscreen tasks are described below. As the Find the Stars task was new, there was initial task
development and piloting was done with the following configurations 2, 3, and 4 items followed by 5, 6, and 7 items where performance was at ceiling (n=11). To determine appropriate load levels for preschoolers and refine the task, the following three configurations presented to children: 8, 9, and 10 items (n=13), 11, 12, and 13 items (n=22), and 9, 11, and 13 items (n=31). Of the 69 tested on the final battery, the majority of the sample completed all 3 tasks (n=56, 81%), with 19% (n=13) who did not complete all tasks.

*Find the Stars task.*

**Practice phase.** In the practice phase the experimenter said “Do you think I can find a hidden star?” followed by a demonstration of how to retrieve a star from one square on a screen by pressing it. This resulted in the presentation of an exciting video of “Jumping Man.” Then a new trial was presented to the child and the experimenter asked “Do you think you can find a hidden star, too?” Then the child was given a chance to retrieve the star to find “Jumping Man.” Following successful retrieval, the child also received the “Jumping Man” video.

**Test phase.** During the test phase the experimenter asked the child, “Do you think you can find even more hidden stars?” then they were presented with a series of squares, each with a hidden star to be found (see Figure 12). Children did not find “Jumping Man” until all stars were found. If children made a predetermined number of search errors (6 greater than the number of items on the screen) prior to finding all the stars the trial ended and a new one started. Children received 3 trials of 3 different loads for a total of 9 test trials.
Figure 12. *Find the Stars* trial configuration of varying loads and retrieval phase where stars were found.

**Motor-Spatial Touchscreen task.**

**Demonstration phase.** The experimenter faced the child and said, “Watch me!” and then proceeded to touch pictures in the target sequence (e.g. A→B→C) (see Figure 4). Having touched all three pictures in the correct order after the first demonstration, the computer played “Jumping Man” accompanied by music. While “Jumping Man” was playing the experimenter turned back to the child, smiled, clapped and said “Yay, I found Jumping Man!” The experimenter repeated this demonstration procedure for two more trials (total of three demonstrations).

**Test phase.** Following the third demonstration, the experimenter faced the child and explained “Yay! I found Jumping Man! Ok, now it’s your turn. Can you find Jumping Man?” If the child imitated the correct sequence they were tested up to two more times on the same configuration (see Figure 3 & 4). Following successful completion of 3-item sequence (3 correct test trials) the
procedure was repeated for a 4-item sequence and following successful completion of a 4-item sequence the procedure was repeated for a 5-item sequence (see Figure 3 & 4). If the child was having difficulty the experimenter reminded them “Remember, start with picture number 1.” If they did not imitate the correct sequence 3 times after 6 failed attempts the task ended and did not progress to a higher level.

Figure 13. Motor-Spatial Task Loads: Three identical pictures appear simultaneously on the screen. On each trial their spatial configuration remains the same, while the identity varies. Based on performance pictures varied in difficulty from 3- to 5-items.

Figure 14. Motor-Spatial Task Test: Children respond to target locations on the screen (as demonstrated).
Flanker Fish Task. **Practice phase:** In the practice phase the experimenter holds the touchscreen and asks the child a question that varies depending on the stimuli presented for 8 trials (See Figure 5). When the child presses a correct arrow they receive a “woo hoo” sound and an incorrect error press is followed by a buzz sound. A correct response advances to the next trial and an incorrect trial presents the same test trial again.

**Test phase:** Following successful completion of the practice phase children proceeded to the test phase. To help the child transition the experimenter reminded them that “For this part of the game we are going to watch the center fish and press the arrow buttons.” This is because in the practice phase some of the trials included single fish or pressing the center fish. The test was comprised of 40 test trials (half incongruent and half where the fish pointed left) and the order of the trials were randomized by participant (Figure 6). If a child did not respond within 5000 ms a fixation star then appeared and a new trial began.

*Figure 15. Flanker Fish sample practice trials: 8 trials are performed with verbal prompts from the experimenter to ensure child understands the task before moving on to critical test trials.*
Figure 16. Flanker fish test trials: 40 test trials are conducted. Half were incongruent, meaning the fish faced a different way from the other fish and half were facing left. Trials were randomized for each participant.

Coding

Find the Stars task: The touchscreen automatically recorded the sequence of behaviors for each trial on the VSWM task.

Find the Stars Span. From this working memory span was computed, or the number of correct presses prior to an error averaged across three tries, for each load. The maximum span score is equivalent to the load, or number of items per trial.

Find the Stars Composite. Each child completed 3 different loads and a composite score was calculated by summing these three working memory spans. Due to slight variations in the three loads tested the maximum composite score ranged from 27 to 36, based on the sum of the three loads presented to the child.

Motor-Spatial Touchscreen task: The touchscreen automatically recorded the sequence of behaviors for each trial on the motor-spatial task.
**Motor-Spatial Ratio Score.** From this a ratio score of the number of correct presses over the number of trials to complete the sequence was computed. This coding approach corresponds with those used in previous publications using this paradigm (e.g. Subiaul, et al., 2015).

**Motor-Spatial Composite Score.** Each child completed 3 different item lengths and a composite score was calculated by summing these three ratio scores. The maximum composite score is 12, based on the sum of the three item lengths presented to the child. One outlier score of .83 was removed from the final analyses as it was greater than two standard deviations below the mean ($M=6.66$, $SD=2.80$).

**Flanker Fish task:** The touchscreen automatically recorded the presses for each trial on the inhibitory control task.

**Flanker Fish Accuracy score.** The primary score computed is accuracy which is the total number of correct responses out of 40.

**Flanker Fish Reaction Time scores.** Two additional scores are computed, average reaction time for congruent trials and average reaction time for incongruent trials.

**Flanker Fish Difference score.** This is computed by subtracting the mean reaction time on correct congruent trials from correct incongruent trials.

### Results

**Data analysis plan**

Each task (Find the Stars, Flanker task, Motor-spatial imitation task) will first be analyzed to assess the effects of cognitive load, age of the child and sex of the child. First omnibus analyses will be reported for each task using repeated measures analysis and assessing the effects of age as a covariate. Given that this is the development phase of the Find the Stars task, however, the data were not collected for each cognitive load. This may have led the design
to be underpowered. For this reason, follow-up 1-way ANOVAs were conducted using age in years as the age variable. Although this potentially increases the risk of type II error, it increases the power of the design to detect age-related differences. Then composite scores for each task were calculated a MANOVA comparing performance on each task was conducted again using age in years to increase the power of the design. Finally, the relationship between the different tasks was analyzed using first order correlations to assess the potential impact of demographic factors and the relationship to the parent-report working memory and inhibitory control scales of the BRIEF-P. Finally, a regression analysis was conducted to assess whether age of the child, performance on the inhibitory control flanker task or performance on the motor-spatial imitation task predicted performance on the newly developed Find the Stars composite score.

**Age and Cognitive Load Effects: Find the Stars**

A subset of participants completed 9, 11, and 13 item trials. For this group, a repeated measures ANOVA with a Greenhouse-Geisser correction determined that there was a main effect of working memory span on performance, $F(2,28) = 4.83, p<.05$, no main effect of age in months $F(2, 28) < 1, p = ns$, no main effect of sex $F(2, 28) < 1, p = ns$, no interaction between working memory span and age in months $F(2, 28) < 1, p = ns$, and no interaction between working memory span and sex $F(2, 28) < 1, p = ns$. These results indicate working memory impacted performance but it did not differ by age in months or sex.

Due to the fact that this was a new task, exploratory follow-up ANOVAs were conducted with the portion of the sample that completed the 9, 11, and 13 sequence using age in years. To examine the effect of age (3, 4, and 5 years) on VSWM a one-way ANOVA on mean working memory span scores at 9-, 11-, and 13-items were run. Scores for 9-items yielded no significant main effect of age in years $F(2, 45) = 2.34, p=ns$. Scores for 10-items yielded no significant main
effect of age in years $F(2, 17) < 1$. Scores for 11-items yielded a significant main effect of age in years $F(2, 53) = 3.91, p < .05$, $\eta^2 = .14$ with 3-year-olds performing significantly worse than 4- and 5-year-olds. Scores for 12-items yielded no significant main effect of age in years $F(2, 21) = 1.70, p = ns$. Scores for 13-items yielded a significant main effect of age in years $F(2, 49) = 9.17, p < .001$, $\eta^2 = .29$ with 3-year-olds significantly worse than 4- and 5-year-olds.

Table 9. Participants by Find the Stars Item Length

<table>
<thead>
<tr>
<th>Measure</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. FTS mean span 9</td>
<td>17</td>
<td>20</td>
<td>8</td>
<td>35</td>
</tr>
<tr>
<td>2. FTS mean span 10</td>
<td>7</td>
<td>9</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>3. FTS mean span 11</td>
<td>17</td>
<td>23</td>
<td>13</td>
<td>53</td>
</tr>
<tr>
<td>4. FTS mean span 12</td>
<td>5</td>
<td>10</td>
<td>6</td>
<td>21</td>
</tr>
<tr>
<td>5. FTS mean span 13</td>
<td>15</td>
<td>21</td>
<td>13</td>
<td>49</td>
</tr>
</tbody>
</table>
Figure 17. Mean span scores by cognitive load and composite score on Find the Stars by age group

**Flanker Fish**

Mean reaction time (RT) for correct congruent trials decreased by age ($M_3 = 2226.55$ ms, $SD = 397.27$; $M_4 = 2212.11$ ms, $SD = 394.79$; $M_5 = 1905.54$ ms, $SD = 503.41$) and correct incongruent trials ($M_3 = 2399.13$ ms, $SD = 527.11$; $M_4 = 2376.78$ ms, $SD = 558.70$; $M_5 = 2063.95$ ms, $SD = 508.40$) decreased by age. Flanker difference scores (incongruent RT - congruent RT) decreased by age ($M_3 = 172.58$ ms, $SD = 434.13$; $M_4 = 164.67$ ms, $SD = 544.92$; $M_5 = 158.42$ ms, $SD = 206.12$), however there was significant variation across participants as evidenced by the large standard deviations. The wide variation in RT difference scores makes this measure uninterpretable and thus we take the approach of examining reaction times and accuracy scores separately as reported in previous studies (Frizzell & Courage, 2015, Zelazo et al., 2013).
A 2(sex; male, female) repeated measures ANOVA on accuracy (congruent, incongruent trials) with Greenhouse-Geisser correction with age in months as a covariate determined that there was a main effect of accuracy on congruent and incongruent trials $F(1, 61) = 36.81, p < .001$, no main effect of age in months $F(2, 61) < 1, p = ns$, no main effect of sex $F(1, 61) < 1, p = ns$, and no interaction between accuracy and age $F(1, 61) < 1, p = ns$, or accuracy and sex $F(1, 61) < 1, p = ns$. These results indicate a significant difference between incongruent and congruent accuracy with more accurate performance on the congruent trials, but this does not vary by age in months or sex of the child. There was however, a trend for an interaction with age and sex of the child ($p = .07$) indicating that older males were more likely to be more accurate on congruent trials than females but there was no difference between younger males and females congruent trials or across ages with males and females on incongruent trials.

A follow up analysis using age in years was conducted and is consistent with prior reporting of results with preschoolers (Frizzell & Courage, 2015). To examine the effect of age on inhibitory control a one-way ANOVA on Flanker Fish accuracy scores yielded a significant main effect of age $F(2, 62) = 15.99, p < .001, \eta^2 = .35$. Post-hoc SNK tests showed that with 3-, 4-, and 5-year-olds were significantly different from one another.

A 2(sex; male, female) repeated measures ANOVA on RT during congruent and incongruent trials with a Greenhouse-Geisser correction determined that there was a main effect of congruency of trials on RT $F(1, 61) = 4.84, p < .05$, no main effect of age in months $F(1, 70) = 1.44, p = ns$, no main effect of sex $F(1, 61) < 1, p = ns$, no interaction between reaction time and age in months $F(1, 61) < 1, p = ns$, and a significant interaction between reaction time and sex $F(1, 61) = 3.6, p < .05$. This interaction was due to a faster RT by males on the incongruent trials and no difference in RT between males and females on the congruent trials.
Figure 18. Mean accuracy scores on Flanker Fish by age in years

Motor-Spatial Imitation

To examine the effect of age on spatial imitation a repeated measures ANOVA with a Greenhouse-Geisser correction on the cognitive load (3, 4, and 5 item ratio scores) across age and sex of the child determined that there was a main effect of cognitive load on performance, $F(2, 58) = 17.76, p < .001$, no main effect of age in months $F(2, 58) < 1, p = ns$, and no main effect of sex $F(1, 58) < 1, p = ns$. There was no interaction between cognitive load and sex $F(2, 58) < 1, p = ns$, and there was no interaction between cognitive load and age in months, $F(2, 58) < 1, p = ns$. These results indicate cognitive load impacted performance but it did not differ by age in months or sex.

The sample size is small and while the age in months variable is well distributed across the sample, a series of exploratory ANOVAs were conducted to examine whether there might in fact be age-related changes by years. A one-way ANOVA on mean ratio scores at 3-, 4-, and 5-items were run. Scores for 3-items yielded no trend for a main effect of age in years $F(2, 78) =
Scores for 4-items yielded a significant main effect of age in years $F(2, 73) = 8.32, p = .001, \eta^2 = .19$. Post-hoc SNK tests showed that with 3-and 4-year-olds performed significantly worse than 5-year-olds. Scores for 5-items yielded a significant main effect of age in years $F(2, 62) = 11.76, p < .001, \eta^2 = .29$, post-hoc SNK ($p < .05$) showed that 3-year-olds performed significantly worse than 4- and 5-year-olds. The 5-item data must be interpreted cautiously however because of the very poor performance of 3-year-olds on the task.

Figure 19. Mean ratio scores by cognitive load and composite score on Motor-Spatial Imitation by age group

Comparing performance on the three tasks.
A 2(sex) x 3(age in years) MANOVA on the composite scores of the three measures (Find the Stars, Flanker task, Motor-spatial imitation), indicated that there was a main effect of age on the Find the Stars task, $F(2, 58) = 8.47, p = .006$ on the Flanker task, $F(2, 58) = 12.65, p = .001$, and on the Motor-Spatial Imitation task, $F(2, 58) = 26.70, p < .001$. There was a main effect of sex on
Motor-Spatial Imitation, $F(2, 58) = 27.70, \ p< .05$, and an interaction between age and sex on the Motor-Spatial Imitation task $F(2, 58) = 8.79, \ p= .01$, but not on any of the other tasks. The interaction indicated that males had higher ratio scores at 4 and 5 years than females but there was the reverse pattern was true at 3 years. Overall, this indicates that these composite measures provide a robust assessment of the performance on the three tasks but there are age-related changes in performance on these tasks.

**Relationship between Find the stars, Motor-Spatial Imitation and Flanker Fish tasks, demographics and parent report**

In order to compare task performance to other variables the composite scores for each task were correlated with demographics and parent report measures. Given that there were age related changes on each task, age was controlled. Partial correlations controlling for age for each composite score, demographics and selected parent report subscore and composite scores on the BRIEF-P are presented in Table 10.

**Task relationships.** There were significant associations between *Find the Stars* and *Flanker* $r(34) = .41, \ p< .01$, and *Flanker* and *Motor-spatial imitation* $r(34) = .47, \ p< .01$, and a trend for *Find the Stars* and *Motor-spatial imitation* $r(34) = .34, \ p = .05$.

**Parent-report relationships.** Due to the reverse coding nature of the BRIEF-P a negative association should be interpreted as a direct correlation. For *Find the Stars*, there were significant correlations with the working memory subscale $r(34) = -.34, \ p < .05$ and EMI composite score of working memory + plan/organize $r(34) = -.33, \ p < .05$. A similar pattern was seen for *Flanker Fish*; there were significant correlations with the working memory subscale $r(34) = -.37, \ p < .05$, plan/organize subscale $r(34)=-.34, \ p<.05$, and EMI composite score of working memory +
plan/organize $r(34) = -0.40$, $p < .05$. For *Motor-spatial imitation*, there was a trend for a significant correlation with the inhibit subscale $r(34) = 0.37$, $p = 0.09$, suggesting that worse inhibit scores reported by parents were associated with better performance on Motor-Spatial Imitation. Negative correlations were as expected and to rule out other possible explanations it was determined that there were no outliers on the inhibit subscale. This trend for a positive correlation between the Inhibit subscale and Motor-Spatial Imitation was surprising. One possible explanation is related to the fact that the inhibit subscale questions are very different from the inhibition measured on these tasks (through perseveration). Another important thing to note is that the inhibit subscale does not correlate with the flanker task as well.
Table 10. Partial Correlation Table for Find the Stars (FTS) composite, Motor-Spatial (MS) composite, Flanker Fish (Flanker), BRIEF-P, and demographic variables controlling for age in days

<table>
<thead>
<tr>
<th>Measure</th>
<th>FTS composite</th>
<th>MS composite</th>
<th>Flanker accuracy</th>
<th>Parental Education</th>
<th>Bilingual</th>
<th>Touchscreen</th>
<th>BRIEF WM</th>
<th>BRIEF Inhibit</th>
<th>BRIEF Plan</th>
<th>BRIEF EMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS</td>
<td>.33†</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Flanker</td>
<td>.53**</td>
<td>.55**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Education</td>
<td>.02</td>
<td>-.21</td>
<td>-.09</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bilingual</td>
<td>-.32†</td>
<td>-.30†</td>
<td>-.32†</td>
<td>-.05</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Touchscreen</td>
<td>.007</td>
<td>.29†</td>
<td>-.2</td>
<td>-.26</td>
<td>-.24</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BRIEF WM</td>
<td>-.34*</td>
<td>.02</td>
<td>-.42*</td>
<td>.06</td>
<td>.22</td>
<td>-.06</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BRIEF Inhibit</td>
<td>-.12</td>
<td>.32†</td>
<td>-.02</td>
<td>.06</td>
<td>.17</td>
<td>-.15</td>
<td>.66**</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BRIEF Plan</td>
<td>-.27</td>
<td>.09</td>
<td>-.34*</td>
<td>.07</td>
<td>.29†</td>
<td>-.06</td>
<td>.73**</td>
<td>.67**</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BRIEF EMI</td>
<td>-.33*</td>
<td>.02</td>
<td>-.40*</td>
<td>.07</td>
<td>.27</td>
<td>-.09</td>
<td>.97**</td>
<td>.73**</td>
<td>.88**</td>
<td>-</td>
</tr>
</tbody>
</table>

\(p < .10, \star p < .05, \star\star p < .01\)
A linear regression with items simultaneously entered was conducted to test if Flanker Fish, Motor-spatial imitation, or age in days significantly predicted VSWM on Find the Stars. The results of the regression indicated that only Flanker Fish significantly predicted Find the Stars scores $\beta = .34, p < .05$, and it explained $23\%$ of the variance $R^2 = .23, F(3, 56) = 5.27, p < .01$ (Table 1). Next, an omnibus MANOVA was conducted to test the age (3, 4, and 5) and sex (male female) effects on all three tasks. The results of the MANOVA indicated there was a main effect of age in years (3, 4, and 5) for Find the Stars $F(1, 49) = 8.47, p < .001$, Flanker Fish $F(1, 49) = 12.65, p = .001$, and Motor-Spatial Imitation $F(1, 49) = 26.70, p = .006$. There was a main effect of sex for Motor-Spatial Imitation $F(1, 49) = 6.12, p < .05$, but not Find the Stars $F(1, 49) < 1, p = ns$ or Flanker $F(1, 49) < 1, p = ns$. Lastly, there was an interaction between age and sex on Motor-Spatial Imitation $F(1, 49) = 8.79, p = .005$ but not Find the Stars $F(1, 49) < 1, p = ns$ or Flanker $F(1, 49) < 1, p = ns$. These findings indicate that performance on all three tasks significantly improved with each additional year, with sex differences and an interaction between sex and age emerging for only Motor-Spatial imitation. Thus older boys performed significantly better on this task than older girls.
Table 11. Predictors of *Find the Stars*

| Variable               | Unstandardized Coefficients | Standardized Coefficients |
|------------------------|-----------------------------|****************************|
|                        | B  | Std. Error | β  | t  | Sig. |
| Constant               | 15.27 | 3.99     | 3.82 | .00 |
| Age in Months          | .05  | .07      | .10 | .80 | .43  |
| Flanker Accuracy       | .32  | .10      | .51** | 3.36 | .002 |
| Motor-Spatial Composite| .04  | .27      | .02 | .14 | .89  |

\[ R^2 = .54 \]

\[ Adjusted\ R^2 = .24 \]

\[ F = 5.95 \]

**p<.001

Discussion

The present study reports on the development of a new VWM task for preschoolers. The study demonstrated that children were able to complete the *Find the Stars* task and that there were age-related changes on the task as a function of increasing cognitive load. These findings suggest that *Find the Stars* may be a more developmentally appropriate task for preschoolers than the current non-verbal WM tasks (e.g. CANTAB SWM). Similar age-related changes were observed on the other two tasks, the flanker task and motor spatial imitation task.
The findings also offer a greater understanding of the relationship between VSWM, motor-spatial imitation, and inhibitory control in 3- to 5-year-olds. There were significant associations on performance between *Find the Stars*, *Flanker Fish*, and *Motor-Spatial Imitation* tasks and with some subscales of parent report BRIEF-P measures. Specifically, *Find the Stars* was positively correlated with the working memory subscale and *Motor-Spatial Imitation* with the inhibit subscale. This is consistent with a recent finding that that BRIEF-P is correlated with a Preschool Executive Function Battery (PEFB; Garon, et al., 2014) in typically developing 3- to 5-year-olds. A regression where age, flanker task performance and motor spatial imitation were simultaneously entered into the model indicated that the *Flanker* task had the most predictive value for performance on the *Find the Stars* task.

This newly developed *Find the Stars task* is promising as the preliminary findings reported here show that children are able to succeed on this task between ages 3 and 5. This is an improvement from the floor performance reported previously on the CANTAB SWM at 4-years. Furthermore, there are age-related changes in success on both *Find the Stars* and Motor-spatial Imitation as a function of cognitive load, or the number of items to keep track of in VSWM. Extending the work of previous studies showing association between cognitive load and motor-spatial imitation (Subiaul & Schilder, 2014) and working memory and inhibitory control (Blakey & Carroll, 2015), the present findings illustrate significant associations across all three tasks.

The age-related changes provide further support for the hierarchical model proposed by Garon and colleagues by documenting the significant improvements on complex working memory and imitation tasks in the preschool years (Garon et al., 2008; 2014). Specifically, there
were both age-related changes on the tasks and age-related differences as a function of cognitive load when varied. It also demonstrates a relationship between the development of VSWM, inhibitory control, and motor-spatial imitation. While the present study is unable to disentangle the complex relationship between working memory and inhibition, the findings suggest that inhibition and attention as measured by the *Flanker Fish* are crucial components of both working memory and motor-spatial imitation. While this does not directly provide support for any of the numerous working memory and EF theoretical models regarding the relationship between inhibitory control to working memory, it suggests that inhibition may be involved in exercising working memory or that they may both depend on the same limited capacity system. Future work to disentangle these components systematically is needed.

One benefit of the current EF battery is that all tasks were presented on a touchscreen and it was quick for children to complete. There are now close to 100,000 educational apps for children in the app store, and approximately 72% of the top apps for sale in Apple’s App Store are in the Toddler/Preschool category. The prevalence of screen media in the lives of young children has increased significantly over the last two decades. The use of touchscreen devices among 2 to 4 year olds in the U.S. increased from 39% to 80% from 2011 to 2013 (Rideout, 2013). However, despite the ubiquity of these apps, our scientific understanding of how children are learning from them is limited. The present study adds to our understanding of how to assess emerging EF abilities using a tool that is both ecologically valid and exciting for preschoolers while also having the benefit of minimizing experimenter error and streamlining the coding process.
While our understanding of EF development is increasing in early childhood, research regarding neural correlates is scarce. The main study that has examined VSWM in preschoolers used functional near infrared spectroscopy (fNIRS), a neuroimaging technique that is well suited for child populations. While being imaged, 3- to 5-year-olds participated in the change detection task (Simmering, 2012) which required children to discriminate between spatial arrays of shape verbally reporting whether the arrays were same or different across trials. Researchers demonstrated that the frontal-parietal network is integral during the change detection task, the network is sensitive to changes in memory load, and parietal cortex activation increased with age (Buss, Fox, Boas, & Spencer, 2014).

Furthermore, comparative neuroscience research indicates that the prefrontal cortex (PFC) is activated in human and non-human adult primates when completing spatial working memory tasks (Damasio & Anderson, 1994; Goldman-Rakic, 1988; Diamond & Goldman-Rakic, 1989; Goldman-Rakic, 1987). Specifically, for the CANTAB SWM task, PET imaging of adults also showed dorsal and ventral prefrontal activation (Owen, Doyon, Petrides, & Evans, 1996).

The present study incorporated an initial piloting and refining of the new Find the Stars task parameters to determine the correct testing protocol for preschoolers. To address the limitation of small sample sizes for the battery of tasks, future work with the final Find the Stars configuration with preschoolers using the 9, 11, 13 item lengths is recommended with additional 3- to 5-year-olds. The Find the Stars task is designed to be versatile such that task complexity may be modified to test with younger infants or older children in the future based on varying the load. It is also designed to be compatible with simultaneous NIRS recording with the ultimate
goal of advancing our understanding of the neural correlates of VSWM and working memory network in early childhood.

Recommendations for future work to understand EF development include using NIRS to examine the neural correlates associated with VSWM, inhibitory control, and motor-spatial imitation in preschoolers due to the fact that EF relies heavily on brain maturation. The present touchscreen tasks in this study have been designed with this ultimate goal in mind. This proposed future direction will strengthen our understanding of normative VSWM development and because of the non-verbal nature of the task and non-invasive NIRS methodology the work will be relevant to those who study clinical populations with VSWM deficits.

This study helps fill a gap in the literature in examining factors associated with VSWM in an age range with existing task limitations. Firstly, it demonstrates a successful, age-appropriate adaptation to the CANTAB SWM by developing and testing *Find the Stars*. These findings contribute to our understanding of the limits of VSWM in preschool. Furthermore, it demonstrates significant relationship between VSWM, inhibitory control, and motor-spatial imitation in preschoolers on both novel and established touchscreen tasks. On both the Find the Stars and Motor-Spatial Imitation tasks there is evidence for an effect of cognitive load and age on performance. Furthermore, parent report on the BRIEF-P EF questionnaire is significantly associated with both new and established EF measures. Lastly, Flanker Fish but not Motor-Spatial Imitation or age in days predicts performance on Find the Stars suggesting the unique role that inhibitory control has within the context of non-verbal VSWM tasks for preschoolers.
CHAPTER V: GENERAL DISCUSSION

Three studies were presented that examined the relationship between task parameters and performance on visuospatial working memory tasks across early childhood. The impact of delay and the predictive nature of productive vocabulary was demonstrated on a simple working memory task (Hide the Pots) at 18 and 24 months. Furthermore, removing perceptual color cues did not disrupt performance at 18 months (Chapter 2). Age-related changes on a complex working memory task (Spin the Pots) were found between 2 and 4 years with 4-year-olds performing significantly better than 2- and 3-year-olds as indexed by a higher success rate and fewer perseverations. Furthermore, there were sex differences with girls performing significantly better than boys as indexed by a higher success rate and fewer perseverations (Chapter 3). Lastly the new Find the Stars VSWM touchscreen task was developmentally appropriate for use with preschoolers and age-related changes were demonstrated for VSWM (Find the Stars), inhibitory control (Flanker Fish), and motor-spatial imitation (Spatial Sequencing Task). Furthermore, higher accuracy on the Find the Stars task was significantly correlated with both Flanker Fish and Motor-Spatial imitation (Chapter 4). These studies help in addressing a gap in our understanding age-related changes in VSWM.

Understanding of working memory development has important implications for early childhood education. Research has linked early working memory abilities in preschool to later math outcomes (Bull, Espy, & Wiebe, 2008). Furthermore, in the school setting, for example, working memory at age 5 is a better predictor of academic success, including reading and math, than IQ (Alloway & Alloway, 2010; Bull & Scerif, 2001). These data to understand typical
VSWM development in early childhood may also inform clinical intervention for children experiencing working memory impairments, such as those with learning disabilities, autism spectrum disorders (ASD), Attention-Deficit Hyperactivity Disorder (ADHD), or acquired brain injury (Stollstoroff, Foss-Feig, Cook Jr., Stein, Gaillard, & Vaidya, 2010).

This conceptualization of age-related changes and factors associated with VSWM performance in early childhood behaviorally is important because it is also during a time period where rapid brain development is occurring. Specifically, by age five children’s brains grow to about 90% of adult size (Debkaban & Sadowsky, 1978). Consistent with EF theories, measuring brain activity can provide additional information about the development of EF through identifying areas specific regions that are recruited during cognitive activity or the relationship among multiple regions recruited that form a neural network.

In the context of VSWM, research indicates that the prefrontal cortex (PFC) is activated in human and non-human adult primates when completing spatial working memory tasks (Beehara, Damasio, Damasio & Anderson, 1994; Goldman-Rakic, 1988; Diamond & Goldman-Rakic, 1989; Goldman-Rakic, 1987). Specifically, for the CANTAB SWM task, PET imaging of adults also showed dorsal and ventral prefrontal activation (Owen, Doyon, Petrides, & Evans, 1996). In 3- to 5-year-olds a change detection task (Simmering, 2012) required children to discriminate between spatial arrays of shape verbally reporting whether the arrays were same or different across trials. Using functional near infrared spectroscopy (fNIRS), researchers demonstrated that the frontal-parietal network is integral during the change detection task, the
network is sensitive to changes in memory load, and parietal cortex activation increased with age (Buss, Fox, Boas, & Spencer, 2014).

By addressing a limitation of the CANTAB SWM (Chapter 4) this dissertation demonstrated the new Find the Stars task is an age appropriate VSWM measure to use with preschoolers. One additional benefit of the task is that it has been designed to be compatible with future NIRS testing. This will allow future work to assess both the behavioral and neural changes taking place during a period of important and rapid developmental change.

There are several hypotheses about how EF components may interact and influence one another. As discussed in the introduction, this dissertation draws on the hierarchical model of EF development (Garon, Bryson, & Smith, 2008) as it is an integrative framework that builds on other existing developmental theoretical frameworks (e.g. Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000; Miyake & Friedman, 2012). This hierarchical model posits that across the first five years of life a developmental progression from early attention to simple skills (e.g. delaying a response) to complex EF processing (e.g. working memory) occurs. Recent work has provided support for this model with a novel EF battery for preschoolers capturing age-related changes in 1.5- to 5-year-olds (Garon, Smith, & Bryson, 2014). Further support for the hierarchical model comes from a recent longitudinal study of 12- to 36-month-olds linking inhibition in infancy to working memory in toddlerhood (Johanson, Marciszko, Brocki, & Bohlin, 2015).

This dissertation adds additional support to the hierarchical model by documenting age-related changes in early childhood on multiple VSWM and EF tasks. Results across experiments
suggest that attention is necessary but not sufficient for successful VSWM performance. In Experiment 1 results that a short delay can disrupt performance suggest that simple working memory skills are present. Additional results in Experiment 2 demonstrate specific support for the model given the age-related changes from 2-to 4-years-olds on the complex VSWM task (Chapter 3). Further support comes from the age-related changes of 3-to 5-year-olds on a battery of tasks (Chapter 4) because as the system becomes more mature as simple working memory, inhibition, and shifting EF components develop. The positive significant relationship between VSWM, inhibitory control, and motor-spatial imitation also suggests that these EF subcomponents are not isolated from one another but integrated or recruiting similar underlying mechanisms which provides additional support for the hierarchical model. This dissertation incorporates both object-based and touchscreen-based tasks, manipulation of cognitive load to examine simple and complex tasks, exploring the relationship to other EF and spatial tasks, and identification of a link between vocabulary and parent report measures of EF. In addition, by examining a broader age-range than is often examined in the literature and manipulating task parameters, it makes an important contribution to our understanding the relationship between cognitive load (task parameters) and age-related changes in VSWM across early childhood.

A current debate in the field is centered around why our memory system has a limited capacity. Early models suggested is was due to a finite number of “slots” that we place representations into (see Chapter 1 for Baddeley and Cowan models). More recently, a resource model of visual memory has been proposed. The main distinction is that there is a limited number of unspecified units that can be stored and that there is a trade-off between the
information stored per item and the number of items (Suchow, Fougnie, Brady, & Avarez, 2014). For example, recruiting working memory to keep track of perceptual color and spatial location information of a few items may require the same resources as remembering many more items with fewer perceptual information. The resource model also addresses an important point—the number of items alone does not entirely determine our ability to remember those items, but that there are other important factors such as color and location. This model takes into account other factors such as the fact that the item configuration can change the accuracy of the representation of those items. Slot models like those of Baddeley and Cowan are not adequate conceptualizations based on the findings from Experiments 1 through 3 in this dissertation.

Moving beyond Baddeley, we see findings from these three experiments relate to the resource models as children are required to track multiple factors to perform successfully on the tasks. In other words, while results indicate children have a limited working memory capacity system, the present findings indicate that multiple factors including timing, directedness, and strategy can impact performance. It is not enough to simply think about children’s working memory as a limited capacity system, but to understand how children are able to track and integrate different cues as their EF system becomes more developed and equipped to handle complex working memory tasks. For example, during Spin the Pots in Experiment 2 (Chapter 3) the color matching strategy spontaneously used by children was associated with better success on the task. This suggests that they may have been able to keep track of fewer items while relying on the retrieved color sticker information to update their memory and flexibly guide their subsequent searches. Another example that is consistent with the resource model comes from
Experiment 3 where differences in the experimenter-directed Motor-Spatial Imitation task and the child-directed newly developed Find the Stars task are present (see Chapter 4).

In integrating the findings across all three experiments we see several common themes emerge. First, small differences in timing can disrupt VSWM performance. For example, increasing the delay from 2 to 10 seconds impacted HTP scores in both 18- and 24-month-olds. Additionally, when the number of items to retrieve increases performance is also poorer such as with STP or higher loads of Find the Stars. Second, strategy during the hiding or retrieval phase may play an important role in VSWM success. This was examined directly in the hiding phase of STP and results indicated children who spontaneously used the color matching strategy had better success on the task. Third, results highlight the importance of considering directedness in the task. HTP in Experiment 1 was experimenter-directed, STP in Experiment 2 was child-directed, and Experiment 3 included both child-directed (Find the Stars) and experimenter-directed (Motor-Spatial Imitation) tasks.

Due to the fact that both task complexity and age varied in Experiment 1 and 2 we cannot determine the role of directedness. However, from the comparison in experiment 3 results indicate that preschoolers perform significantly better on child-directed versus experimenter-directed tasks. Why might children perform better when the task is child-directed? One possible account for this stems from research on the egocentric bias that children have to remember visual information from their perspective rather than from another person’s perspective (Piaget & Inhelder, 1956).
Future studies are needed to systematically examine the role of these three components to further understand the unique role or interaction among various components that can impact performance. One methodology that could be incorporated is eye-tracking to determine how their attention system is encoding the information as well as searching during the test phase. Another future direction is to examine whether or not children in the Find the Stars task who used a linear retrieval strategy (e.g. searching in rows or columns versus sporadically) had better working memory spans. One limitation of developing Find the Stars is that some of the testing involved different span lengths and it was only run with preschoolers. Furthermore, many tasks available for assessing WM are used at targeted ages. For this reason, future work is needed to develop and modify the Find the Stars app parameters (e.g. number of items, locations, directedness, distractor features) to be used with children from 1 to 5. Following this modification, follow up studies to include a more thorough investigation of the role of vocabulary across early childhood is needed. This was demonstrated to predict VSWM in 18- and 24-month-olds on Hide the Pots, but not VSWM in 2- and 3-year-olds on Spin the Pots. By controlling task demands we may gain a clearer sense of this complex relationship and further our understanding of cognitive factors that may underlie the development of VSWM.

Conceptualizing age-related changes in working memory over the course of early childhood is important because it is crucial for complex developmental tasks and problem solving that all require substantial information maintenance, updating, or manipulating of information. It can also very difficult to do given the limitations associated with task comprehension and assessment in early childhood. By studying the development of VSWM
across early childhood, we increase our understanding of the age-related changes working memory and add to our understanding of cognitive development.
REFERENCES


108


